PROBABILITY THEORY & FUZZY LOGIC

Lotfi A. Zadeh

Computer Science Division
Department of EECS
UC Berkeley

URL: http://www-bisc.cs.berkeley.edu
URL: http://zadeh.cs.berkeley.edu/

Email: Zadeh@cs.berkeley.edu

April 24, 2003

BACKDROP

LAZ 4/24/2003

PROBABILITY THEORY AND FUZZY LOGIC

- *How does fuzzy logic relate to probability theory?*
- *This is the question that was raised by Loginov in 1966, shortly after the publication of my first paper on fuzzy sets (1965).*
- *Relationship between probability theory and fuzzy logic has been, and continues to be, an object of controversy.*

PRINCIPAL VIEWS

- *Inevitability of probability*
- *Fuzzy logic is probability theory in disguise*
- *The tools provided by fuzzy logic are not of importance*
- *Probability theory and fuzzy logic are complementary rather than competitive*

CONTINUED

- ***My current view:***

- ***It is a fundamental limitation to base probability theory on bivalent logic***
- ***Probability theory should be based on fuzzy logic***

RELATED PAPER

- Lotfi A. Zadeh, “ Toward a perception-based theory of probabilistic reasoning with imprecise probabilities,” special issue on imprecise probabilities, Journal of Statistical Planning and Inference, Vol. 105, pp.233-264, 2002.

- Downloadable from:

http://www-bisc.cs.berkeley.edu/BISCProgram/Projects.htm

THERE IS A FUNDAMENTAL CONFLICT BETWEEN BIVALENCE AND REALITY

- we live in a world in which almost everything is a matter of degree

but

- in bivalent logic, every proposition is either true or false, with no shades of gray allowed
    - in fuzzy logic, everything is, or is allowed to be, a matter of degree
- in bivalent-logic-based probability theory, PT, only certainty is a matter of degree
    - in perception-based probability theory, PTp, everything is, or is allowed to be, a matter of degree

INEVITABILITY OF PROBABILITY

- *The only satisfactory description of uncertainty is probability. By this I mean that every uncertainty statement must be in the form of a probability; that several uncertainties must be combined using the rules of probability; and that the calculus of probabilities is adequate to handle all situations involving uncertainty…probability is the only sensible description of uncertainty and is adequate for all problems involving uncertainty. All other methods are inadequate… anything that can be done with fuzzy logic, belief functions, upper and lower probabilities, or any other alternative to probability can better be done with probability [Lindley (1987)]*

CONTINUED

- *The numerous schemes for representing and reasoning about uncertainty that have appeared in the AI literature are unnecessary – probability is all that is needed [Cheesman (1985)]*

BASIC PROBLEMS WITH PT

PT

IT IS A FUNDAMENTAL LIMITATION TO BASE PROBABILITY THEORY ON BIVALENT LOGIC

- ***A major shortcoming of bivalent-logic-based probability theory, PT, relates to the inability of PT to operate on perception-based information***
- ***In addition, PT has serious problems with***

(a) brittleness of basic concepts

(b) the "it is possible but not probable" dilemma

PREAMBLE

- *It is a deep-seated tradition in science to strive for the ultimate in rigor and precision. But as we enter into the age of machine intelligence and automated reasoning, other important goals come into view.*

CONTINUED

- We begin to realize that humans have a remarkable capability—a capability which machines do not have—to perform a wide variety of physical and mental tasks without any measurements and any computations. In performing such tasks, humans employ perceptions of distance, speed, direction, size, likelihood, intent and other attributes of physical and mental objects.

CONTINUED

- *To endow machines with this capability, what is needed is a theory in which the objects of computation are, or are allowed to be, perceptions. The aim of the computational theory of perceptions is to serve this purpose—purpose which is not served by existing theories.*

KEY IDEA

- ***In the computational theory of perceptions, perceptions are dealt with through their descriptions in a natural language***

COMPUTATIONAL THEORY OF PERCEPTIONS (CTP) BASIC POSTULATES

- *perceptions are intrinsically imprecise*
- *imprecision of perceptions is a concomitant of the bounded ability of sensory organs—and ultimately the brain—to resolve detail and store information*

KEY POINTS

- *a natural language is, above all, a system for describing and reasoning with perceptions*
- *in large measure, human decisions are perception-based*
- *one of the principal purposes of CWP (Computing with Words and Perceptions) is that of making it possible to construct machines that are capable of operating on perception-based information expressed in a natural language*
- *existing bivalent-logic-based machines do not have this capability*

ILLUSTRATION
AUTOMATION OF DRIVING IN CITY TRAFFIC

- a blind-folded driver could drive in city traffic if
  a) a passenger in the front seat could instruct the driver on what to do
  b) a passenger in the front seat could describe in a natural language his/her perceptions of decision-relevant information
- replacement of the driver by a machine is a much more challenging problem in case (b) than in case (a)

MEASUREMENT-BASED VS. PERCEPTION-BASED INFORMATION

INFORMATION

- **measurement-based** / **numerical**
- **perception-based** / **linguistic**

measurement-based / numerical	perception-based / linguistic
it is 35 C°	It is very warm
Eva is 28	Eva is young
probability is 0.8	probability is high
	it is cloudy
	traffic is heavy
	it is hard to find parking near the campus

- measurement-based information may be viewed as special case of perception-based information

MEASUREMENT-BASED VS. PERCEPTION-BASED CONCEPTS

<u>measurement-based</u>	<u>perception-based</u>
expected value	usual value
stationarity	regularity
continuous	smooth

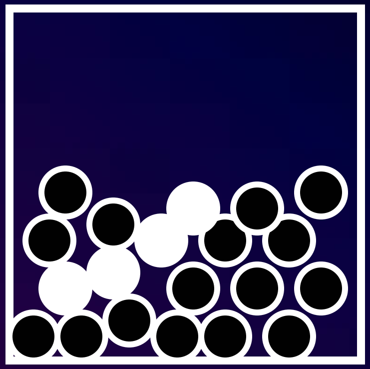
Example of a regular process

$T = (t_0, t_1, t_2 \ldots)$

t_i = travel time from home to office on day i.

WHAT IS CWP?

THE BALLS-IN-BOX PROBLEM



Version 1. Measurement-based

- *a box contains 20 black and white balls*
- *over 70% are black*
- *there are three times as many black balls as white balls*

- *what is the number of white balls?*
- *what is the probability that a ball drawn at random is white?*

CONTINUED

Version 2. Perception-based

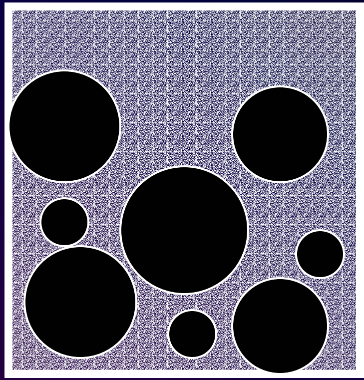
- *a box contains about 20 black and white balls*
- *most are black*
- *there are several times as many black balls as white balls*

- *what is the number of white balls?*
- *what is the probability that a ball drawn at random is white?*

CONTINUED

Version 3. Perception-based

box



- a box contains about 20 black balls of various sizes
- most are large
- there are several times as many large balls as small balls

- what is the number of small balls?
- what is the probability that a ball drawn at random is small?

COMPUTATION (version 1)

- *measurement-based*

  X = number of black balls

  Y_2 number of white balls

  $X \geq 0.7 \cdot 20 = 14$

  $X + Y = 20$

  $X = 3Y$

  $X = 15 \quad ; \quad Y = 5$

  $p = 5/20 = .25$

- *perception-based*

  X = number of black balls

  Y = number of white balls

  $X = most \times 20^*$

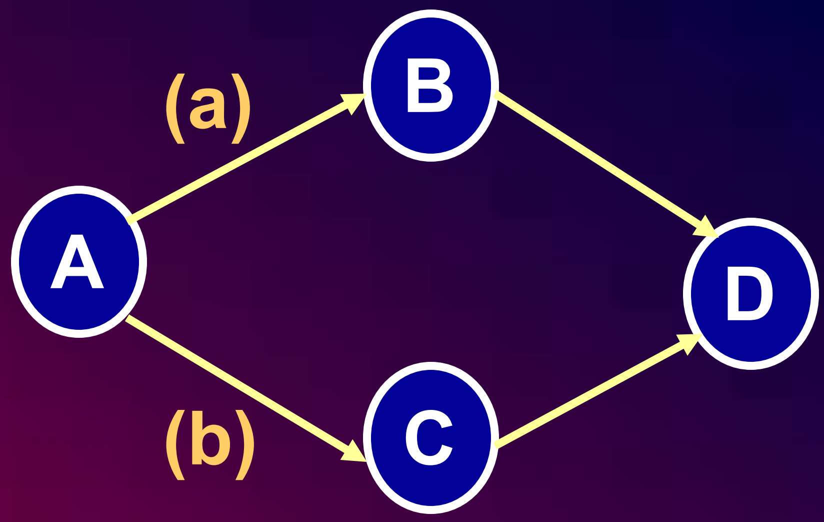
  $X = several\ ^*Y$

  $X + Y = 20^*$

  $P = Y/N$

THE TRIP-PLANNING PROBLEM

- *I have to fly from A to D, and would like to get there as soon as possible*
- *I have two choices: (a) fly to D with a connection in B; or*
  (b) fly to D with a connection in C



- *if I choose (a), I will arrive in D at time t_1*
- *if I choose (b), I will arrive in D at time t_2*
- *t_1 is earlier than t_2*
- *Should I choose (a) ?*

CONTINUED

- *now, let us take a closer look at the problem*
- *the connection time, c_B , in B is short*
- *should I miss the connecting flight from B to D, the next flight will bring me to D at t_3*
- *t_3 is later than t_2*
- *what should I do?*

$$decision = f\,(\,t_1\,,\;t_2\,,\;t_3\,,c_B\,,c_C\,)$$

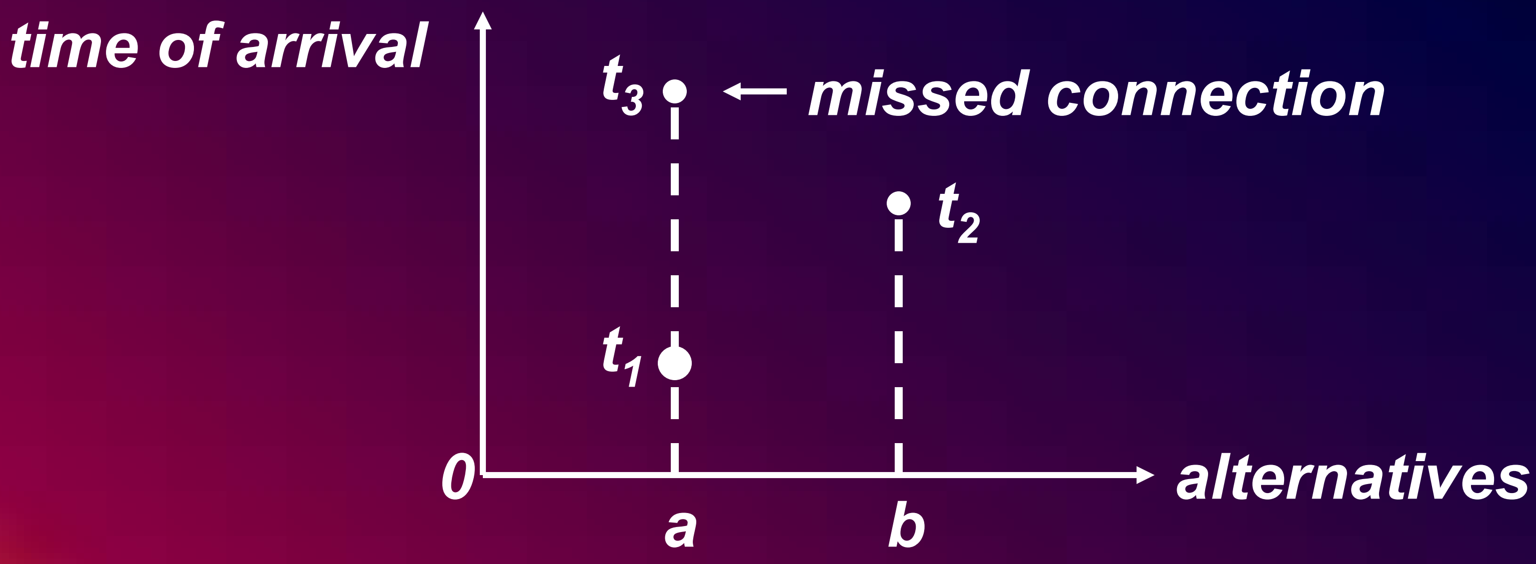
existing methods of decision analysis do not have the capability to compute f

reason: nominal values of decision variables ≠ observed values of decision variables

CONTINUED

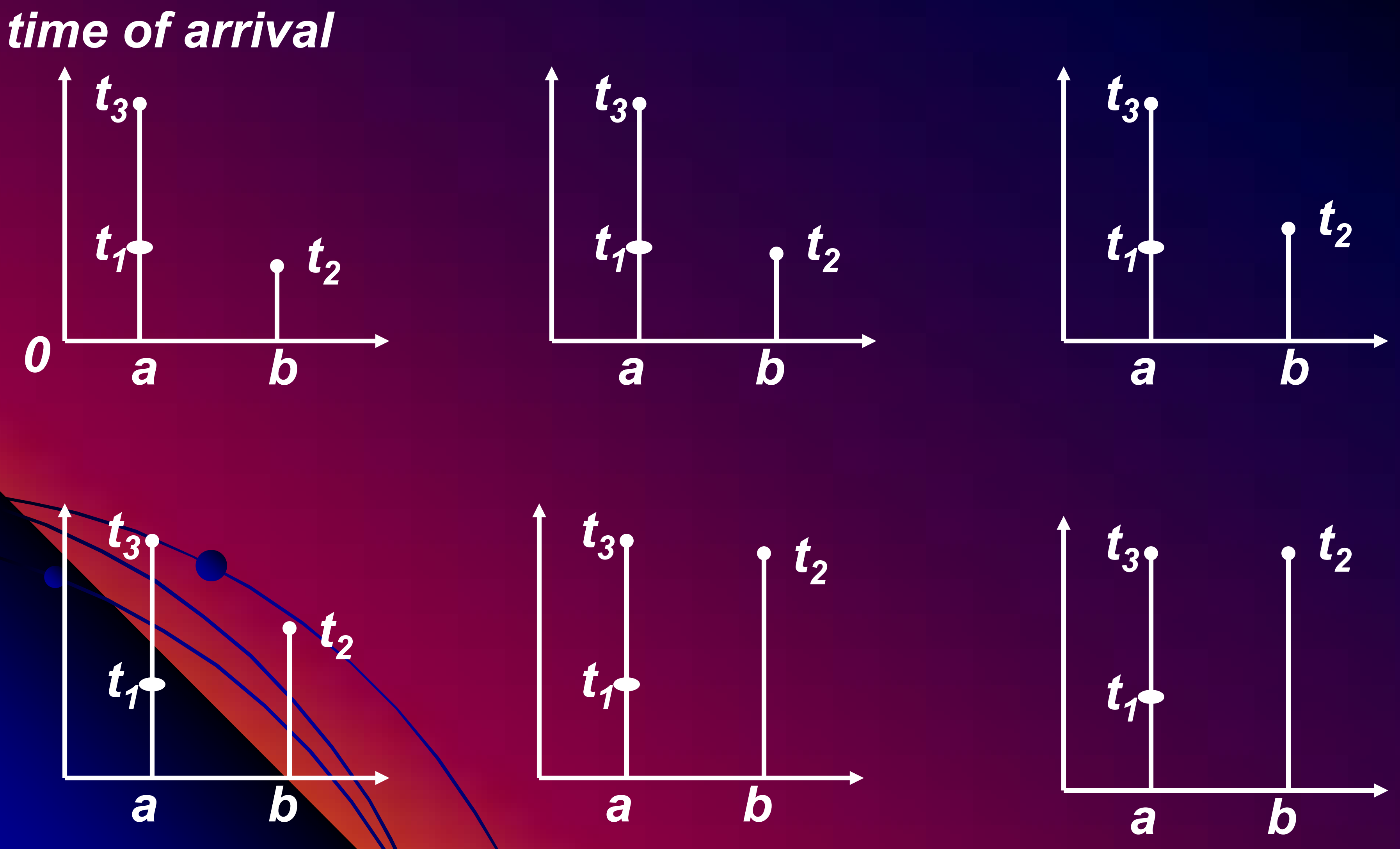
- *the problem is that we need information about the probabilities of missing connections in B and C.*
- *I do not have, and nobody has, measurement-based information about these probabilities*
- *whatever information I have is perception-based*

THE KERNEL PROBLEM —THE SIMPLEST B-HARD DECISION PROBLEM



- decision is a function of t_1, t_2, t_3 and perceived probability of missing connection
- strength of decision

DECISION



LAZ 4/24/2003

TEST PROBLEMS

- *Most Swedes are tall*
- *What is the average height of Swedes?*

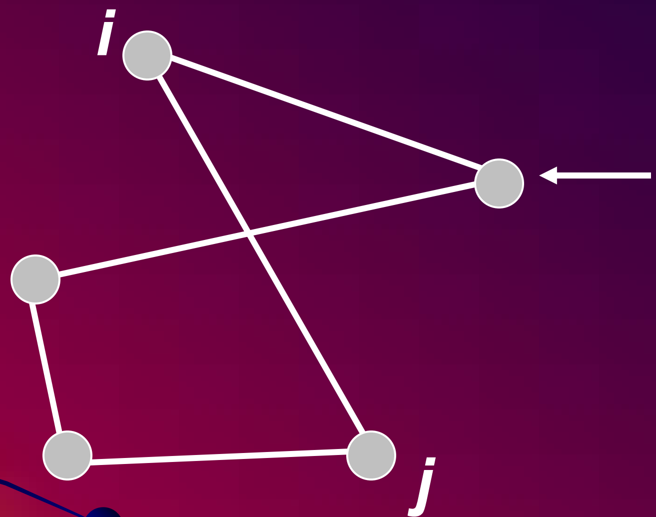
- *Prob {Robert is young} is low*
  Prob {Robert is middle-aged} is high
  Prob {Robert is old} is low

- *What is the probability that Robert is neither young nor old?*

CONTINUED

TSP

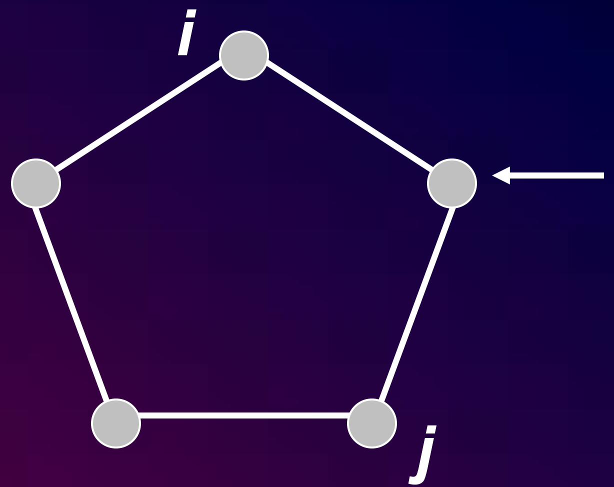
traveling salesman problem



c_{ij} = *measured cost of travel from i to j*

ASP

airport shuttle problem

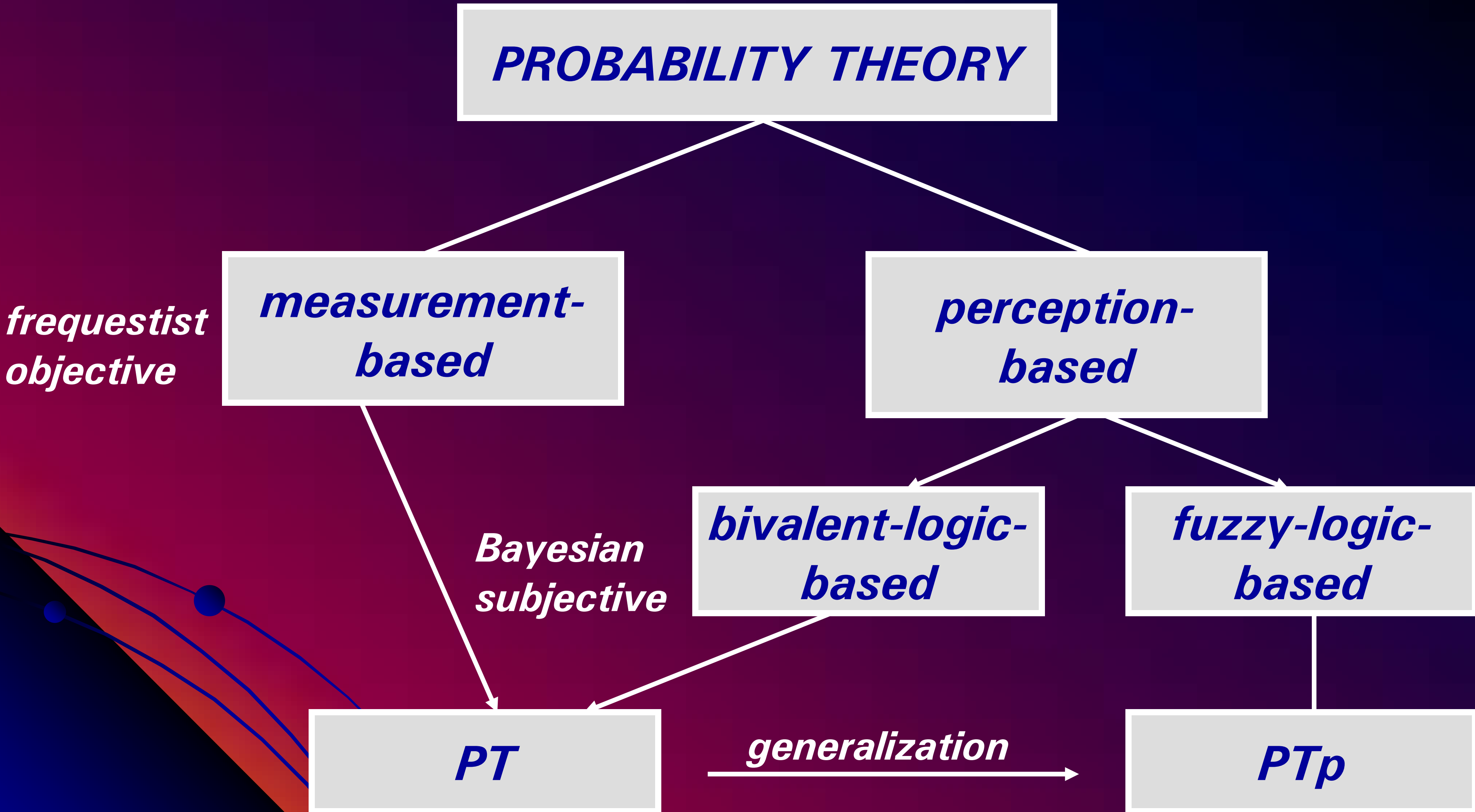


t_{ij} = *perceived time of travel from i to j*

PROBLEMS WITH PT

- Bivalent-logic-based PT is capable of solving complex problems
- But, what is not widely recognized is that PT cannot answer simple questions drawn from everyday experiences
- To deal with such questions, PT must undergo three stages of generalization, leading to perception-based probability theory, PTp

BASIC STRUCTURE OF PROBABILITY THEORY



- *In PTp everything is or is allowed to be perception-based*

THE NEED FOR A RESTRUCTURING OF PROBABILITY THEORY

- *to circumvent the limitations of PT three stages of generalization are required*

1. *f-generalization*
2. *f.g-generalization*
3. *nl-generalization*

$$PT \xrightarrow{\text{f-generalization}} PT^{+} \xrightarrow{\text{f.g-generalization}} PT^{++} \xrightarrow{\text{nl-generalization}} PTp$$

FUNDAMENTAL POINTS

- *the point of departure in perception-based probability theory (PTp) is the postulate:*

subjective probability=perception of likelihood

- *perception of likelihood is similar to perceptions of time, distance, speed, weight, age, taste, mood, resemblance and other attributes of physical and mental objects*
- *perceptions are intrinsically imprecise, reflecting the bounded ability of sensory organs and, ultimately, the brain, to resolve detail and store information*

- *perceptions and subjective probabilities are f-granular*

SIMPLE EXAMPLES OF QUESTIONS WHICH CANNOT BE ANSWERED THROUGH THE USE OF PT

- ***I am driving to the airport. How long will it take me to get there?***

  Hotel clerk: About 20-25 minutes

  PT: Can't tell

- ***I live in Berkeley. I have access to police department and insurance company files. What is the probability that my car may be stolen?***

  PT: Can't tell

- ***I live in the United States. Last year, one percent of tax returns were audited. What is the probability that my tax return will be audited?***

  PT: Can't tell

CONTINUED

- *Robert is a professor. Almost all professors have a Ph.D. degree. What is the probability that Robert has a Ph.D. degree?*

  PT: Can't tell
- *I am talking on the phone to someone I do not know. How old is he?*

  My perception: Young

  PT: Can't tell
- *Almost all A's are B's. Almost all B's are C's. What fraction of A's are C's?*

  PT: Between 0 and 1
- *The balls-in-box example*
- *The trip-planning example*
- *The Robert example*

BRITTLENESS (DISCONTINUITY)

- Almost all concepts in PT are bivalent in the sense that a concept, C, is either true or false, with no partiality of truth allowed. For example, events A and B are either independent or not independent. A process, P, is either stationary or nonstationary, and so on. An example of brittleness is: If all A′s are B′s and all B′s are C′s, then all A′s are C′s; but if almost all A′s are B′s and almost all B′s are C′s, then all that can be said is that proportion of A′s in C′s is between 0 and 1.

BRITTLENESS OF BIVALENT-LOGIC-BASED DEFINITIONS

- *when a concept which is in reality a matter of degree is defined as one which is not, the sorites paradox points to a need for redefinition*
- *stability*
- *statistical independence*
- *stationarity*
- *linearity*
- *…*

BRITTLENESS OF DEFINITIONS

- ***statistical independence***

$$P\,(A, B) = P(A)\,P(B)$$

- ***stationarity***

$$P\,(X_1,\ldots, X_n) = P\,(X_1\text{-}a,\ldots, X_n\text{-}a) \text{ for all } a$$

- ***randomness***

Kolmogorov, Chaitin, …

- ***in PT_p, statistical independence, stationarity, etc are a matter of degree***

BRITTLENESS OF DEFINITIONS (THE SORITES PARADOX)

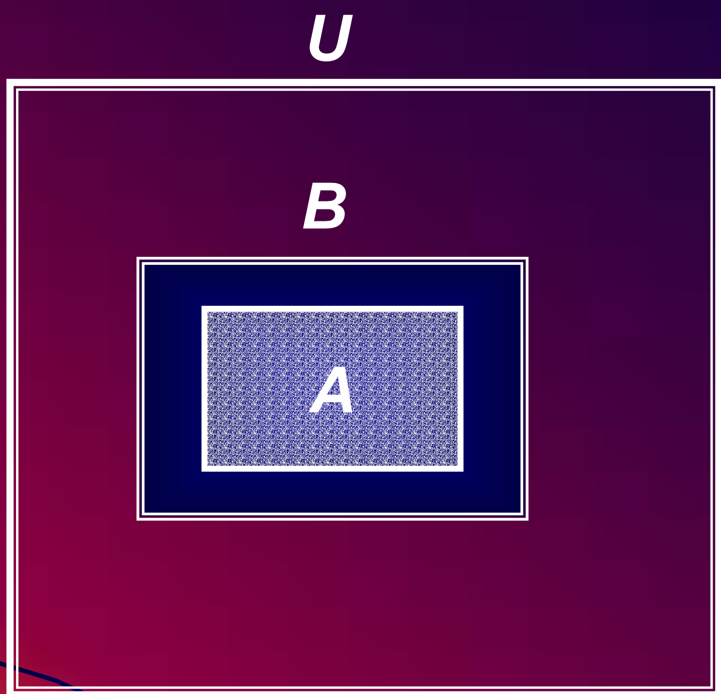
statistical independence

- *A and B are independent* $\longleftrightarrow$ $P_A(B) = P(B)$
- *suppose that (a) $P_A(B)$ and $P(B)$ differ by an epsilon; (b) epsilon increases*
- *at which point will A and B cease to be independent?*
- *statistical independence is a matter of degree*
- *degree of independence is context-dependent*
- *brittleness is a consequence of bivalence*

THE DILEMMA OF "IT IS POSSIBLE BUT NOT PROBABLE"

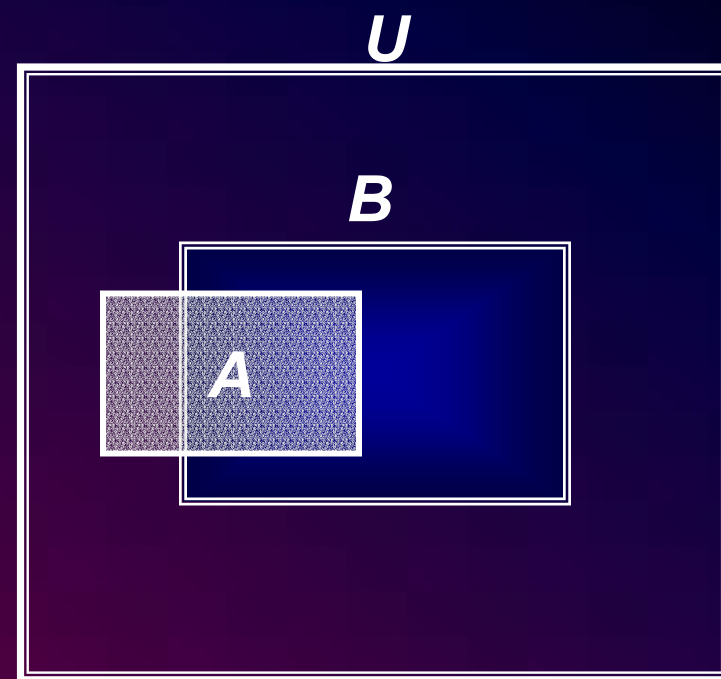
- *A simple version of this dilemma is the following. Assume that A is a proper subset of B and that the Lebesgue measure of A is arbitrarily close to the Lebesgue measure of B. Now, what can be said about the probability measure, P(A), given the probability measure P(B)? The only assertion that can be made is that P(A) lies between 0 and P(B). The uniformativeness of this assessment of P(A) leads to counterintuitive conclusions. For example, suppose that with probability 0.99 Robert returns from work within one minute of 6pm. What is the probability that he is home at 6pm?*

CONTINUED



A= proper subset of B

$$0 \le P(A) \le P(B)$$



$A \cap B$: proper subset of A

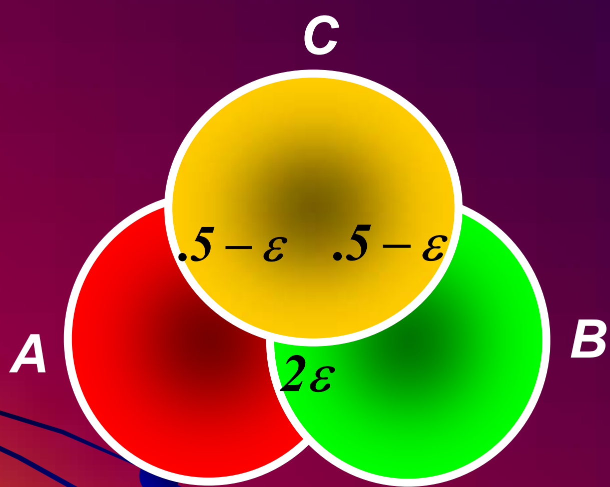
$$0 \le P_A(B) \le 1$$

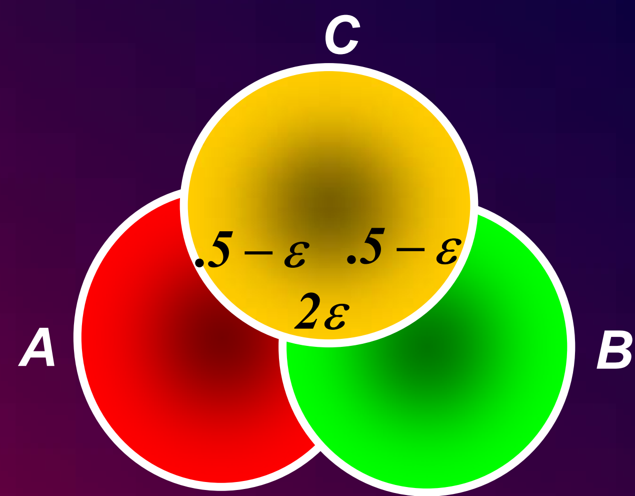
CONTINUED

- ***Using PT, with no additional information or the use of the maximum entropy principle, the answer is: between 0 and 1. This simple example is an instance of a basic problem of what to do when we know what is possible but cannot assess the associated probabilities or probability distributions. A case in point relates to assessment of the probability of a worst case scenario.***

EXAMPLE -- INFORMATION ORTHOGONALITY

- *A,B,C are crisp events*
- *principal dependencies: (a) conjunctive; (b) serial*
- *conjunctive:* $P_{A,B}(C)$=? *given* $P_A(C)$ *and* $P_B(C)$





counterintuitive

$$\mathrm{P}_A(C)=\frac{0.5-\varepsilon}{0.5+\varepsilon}\approx 1$$

$$\mathrm{P}_B(C)=\frac{0.5-\varepsilon}{0.5+\varepsilon}\approx 1$$

$$\mathrm{P}_{A,B}(C)=\frac{0}{2\varepsilon}=0$$

$$\mathrm{P}_A(C)=1$$

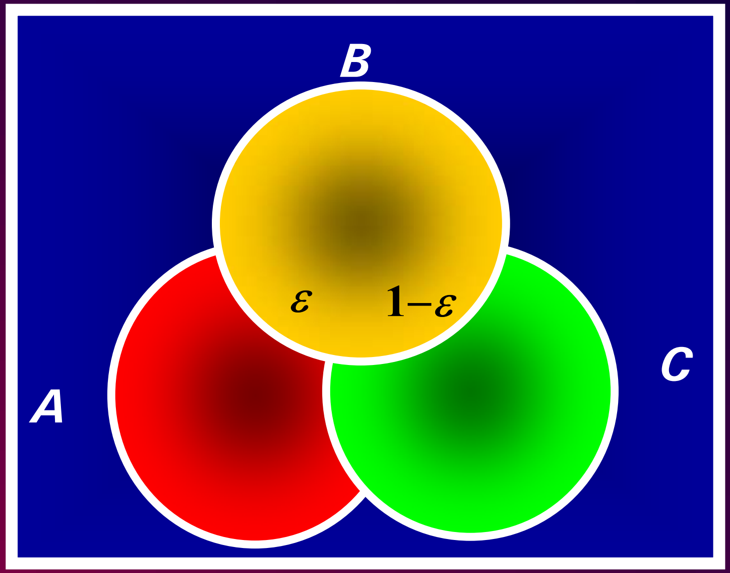
$$\mathrm{P}_B(C)=1$$

$$\mathrm{P}_{A,B}(C)=1$$

SERIAL

$P_A(C) = ?$ given $P_A(B)$ and $P_B(C)$

scenario



counterintuitive

$$P_A(B) = 1$$

$$P_B(C) = 1 - \varepsilon$$

$$P_A(C) = 0$$

REAL-WORLD EXAMPLE

C= US-born

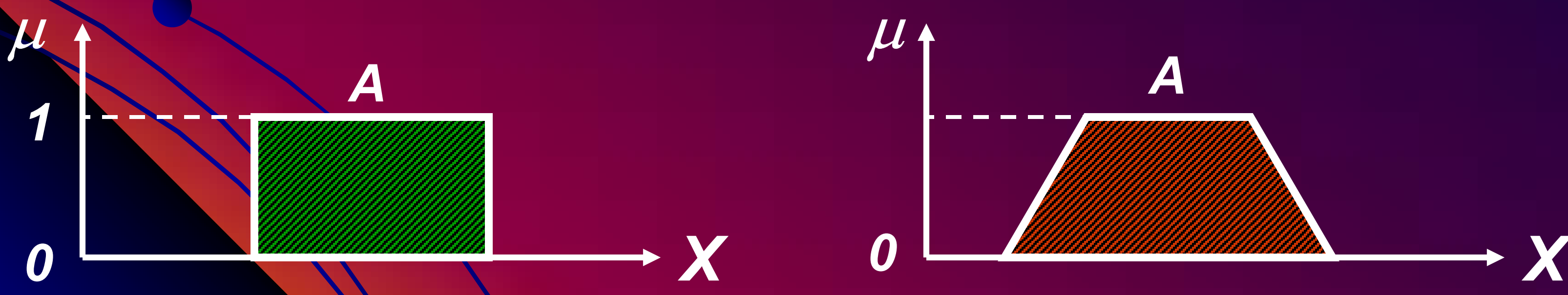
A= professor

B= engineer

- *most engineers are US-born*
- *most professors are US-born*
- *most (engineers^professors) are not US-born*

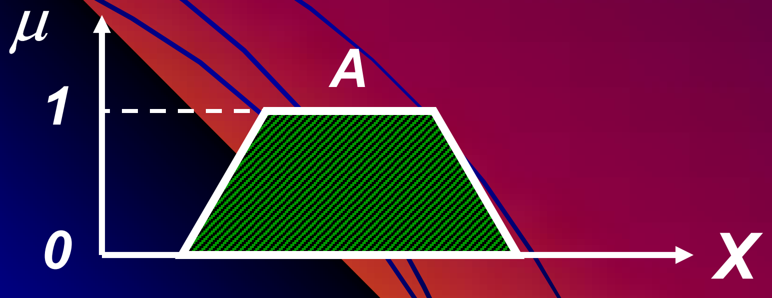
F-GENERALIZATION

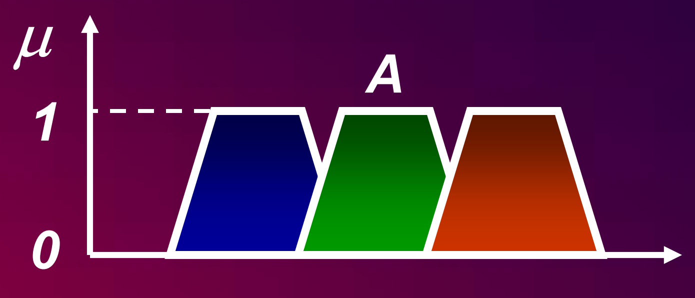
- *f-generalization of a theory, T, involves an introduction into T of the concept of a fuzzy set*
- *f-generalization of PT, PT^+ , adds to PT the capability to deal with fuzzy probabilities, fuzzy probability distributions, fuzzy events, fuzzy functions and fuzzy relations*



F.G-GENERALIZATION

- *f.g-generalization of T, T^{++}, involves an introduction into T of the concept of a granulated fuzzy set*
- *f.g-generalization of PT, PT^{++} , adds to PT^{+} the capability to deal with f-granular probabilities, f-granular probability distributions, f-granular events, f-granular functions and f-granular relations*





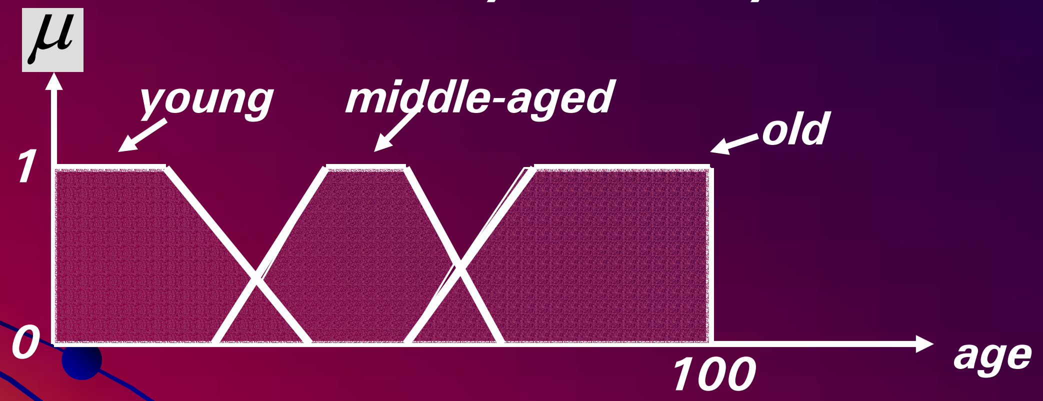
EXAMPLES OF F-GRANULATION (LINGUISTIC VARIABLES)

color: red, blue, green, yellow, …

age: young, middle-aged, old, very old

size: small, big, very big, …

distance: near, far, very, not very far, …



- *humans have a remarkable capability to perform a wide variety of physical and mental tasks, e.g., driving a car in city traffic, without any measurements and any computations*
- *one of the principal aims of CTP is to develop a better understanding of how this capability can be added to machines*

NL-GENERALIZATION

A	*nl-generalization*	***A_p***
crisp set		*PNL-defined set*
PT	*nl-generalization*	***PT_p***

crisp probability → *PNL-defined probability*

crisp relation → *PNL-defined relation*

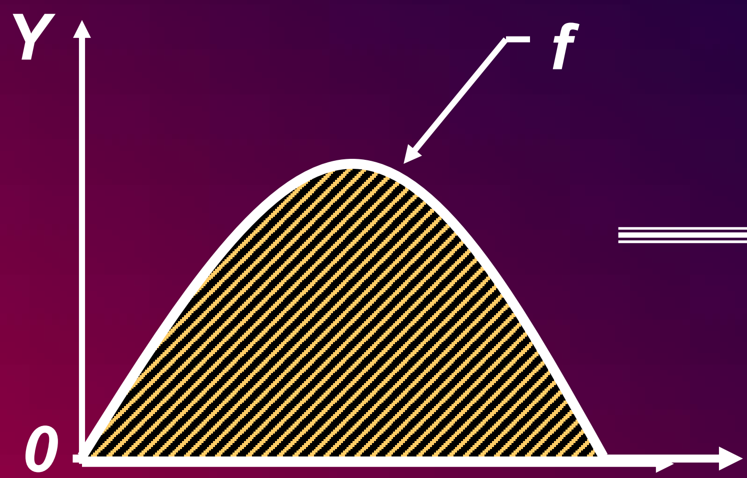
crisp independence → *PNL-defined independence*

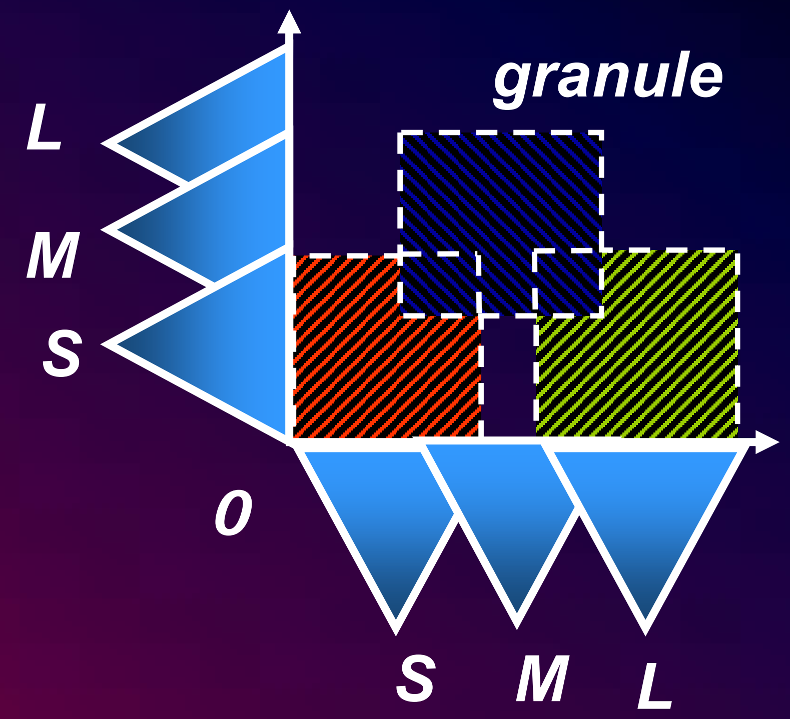
…

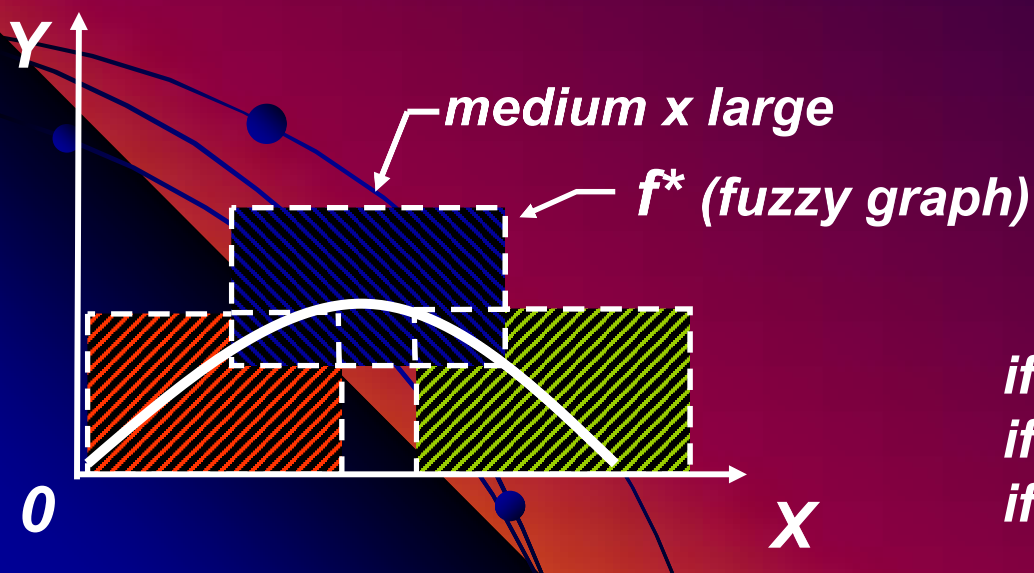
NL-GENERALIZATION

- Nl^{++}-generalization of T. T_{nl}, involves an addition to T^{++} of a capability to operate on propositions expressed in a natural language
- nl-generalization of T adds to T^{++} a capability to operate on perceptions described in a natural language
- nl-generalization of PT, PT_{nl}, adds to PT^{++} a capability to operate on perceptions described in a natural language
- nl-generalization of PT is perception-based probability theory, PTp
- a key concept in PTp is PNL (Precisiated Natural Language)

PERCEPTION OF A FUNCTION







$f \xrightarrow{\text{perception}} f^*$:

if X is small then Y is small
if X is medium then Y is large
if X is large then Y is small

TEST PROBLEM

- *A function, Y=f(X), is defined by its fuzzy graph expressed as*

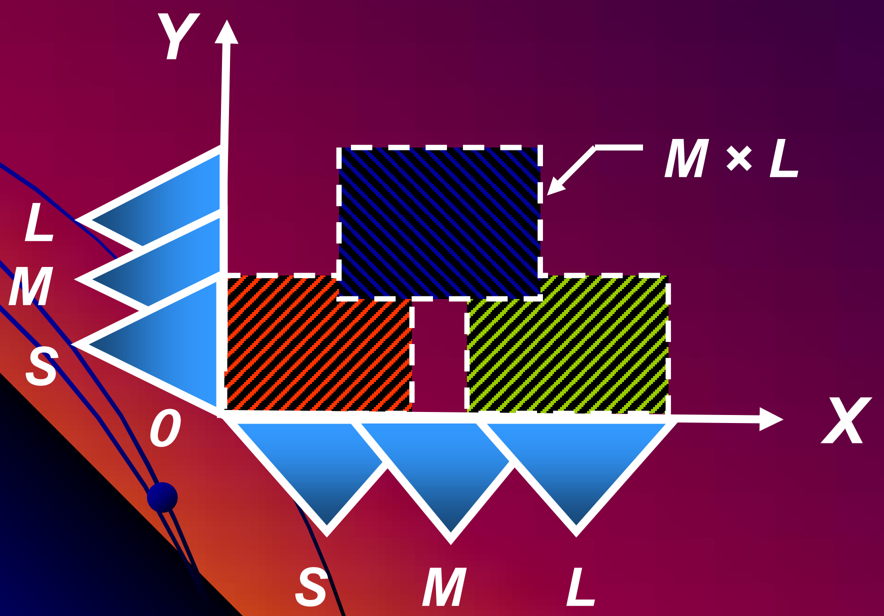
  f_1 if X is small then Y is small

  if X is medium then Y is large

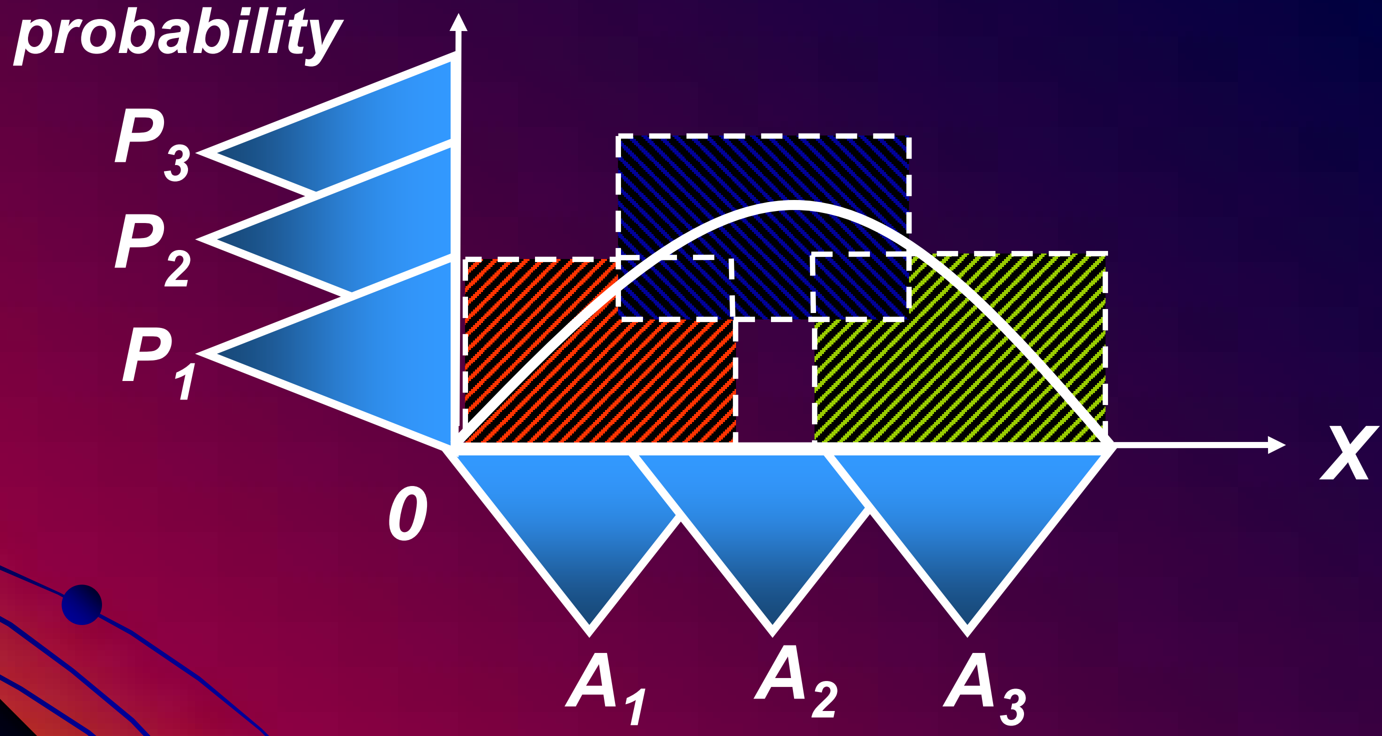
  if X is large then Y is small

  (a) what is the value of Y if X is not large?

  (b) what is the maximum value of Y



BIMODAL DISTRIBUTION (PERCEPTION-BASED PROBABILITY DISTRIBUTION)



$P(X) = P_{i(1)} \backslash A_1 + P_{i(2)} \backslash A_2 + P_{i(3)} \backslash A_3$

Prob $\{X \text{ is } A_i\}$ is $P_{j(i)}$

P(X)= low\small + high\medium + low\large

LAZ 4/24/2003

CONTINUED

- *function: if X is small then Y is large +…*
  (X is small, Y is large)
- *probability distribution: low \ small + low \ medium + high \ large +…*
- *Count \ attribute value distribution: 5* \ small + 8* \ large +…*

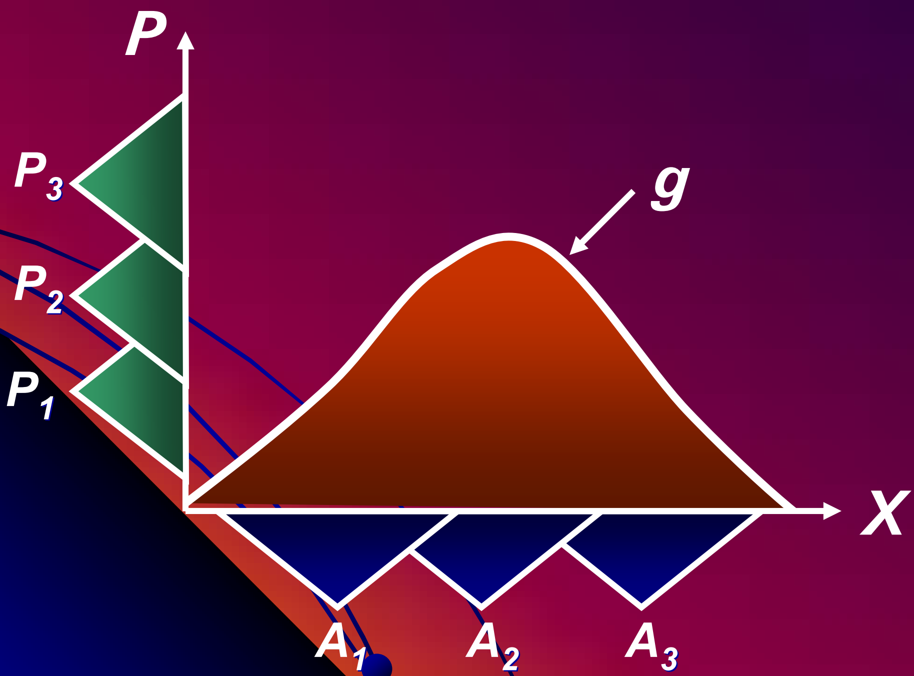
PRINCIPAL RATIONALES FOR F-GRANULATION

- *detail not known*
- *detail not needed*
- *detail not wanted*

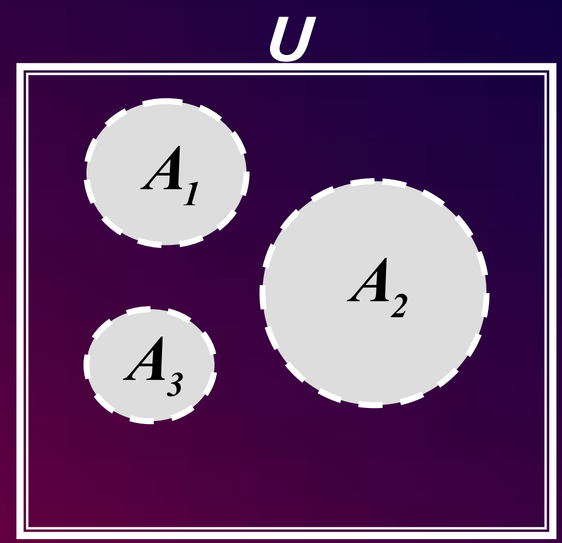
BIMODAL PROBABILITY DISTRIBUTIONS (LAZ 1981)

(a) possibility\probability

(b) probability\\possibility



$$P^* = P_1 \backslash A_1 + \cdots + P_n \backslash A_n$$

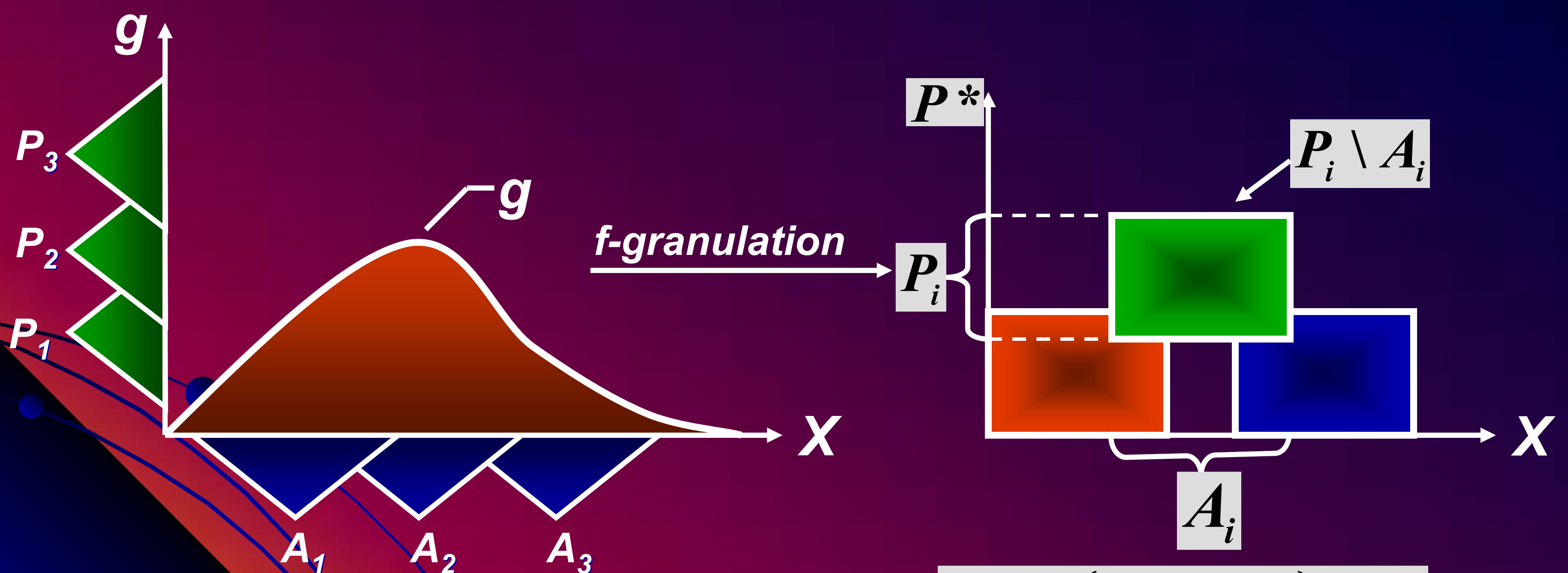


$$P = P_1 \backslash\backslash A_1 + \cdots + P_n \backslash\backslash A_n$$

BIMODAL PROBABILITY DISTRIBUTION

X: a random variable taking values in U

g: probability density function of X



$$P^* = \Sigma_i P_i \backslash A_i$$

$$\text{Prob}\{X \text{ is } A_i\} \text{ is } P_i$$

$$\text{Prob}\{X \text{ is } A_i\} = \int_U \mu_{A_i}(u) g(u) du$$

CONTINUED

P defines a possibility distribution of g*

$$\pi(g) = \mu_{P_i}\left(\int_U \mu_{A_i}(u)g(u)du\right) \wedge \cdots \wedge \mu_{P_u}\left(\int_U \mu_{A_n}(u)g(u)du\right)$$

problems

a) what is the probability of a perception-based event A in U
b) what is the perception-based expected value of X

PROBABILITY OF A PERCEPTION-BASED EVENT

problem: **Prob {X is A} is ?B** **knowing $\pi(g)$**

$$\text{Prob}\,\{\text{X is A}\} = \int_U \mu_A(u)g(u)du = f(g)$$

Extension Principle

$$\frac{\pi_1(g)}{\pi_2(f(g))}$$

$$\pi_2(v) = \sup_g \pi_1(g)$$

subject to: $v = f(g)$

CONTINUED

$$\mu_A(v) = \sup_g (\mu_{P_1}(\int_U \mu_{A_1}(u)g(u)du) \wedge \cdots$$

$$\wedge \mu_{P_n}(\int_U \mu_{A_n}(u)g(u)du))$$

subject to

$$v = \int_U \mu_A(u)g(u)du$$

EXPECTED VALUE OF A BIMODAL PD

$$E(P^*) = \int_U ug(u)du = f(g)$$

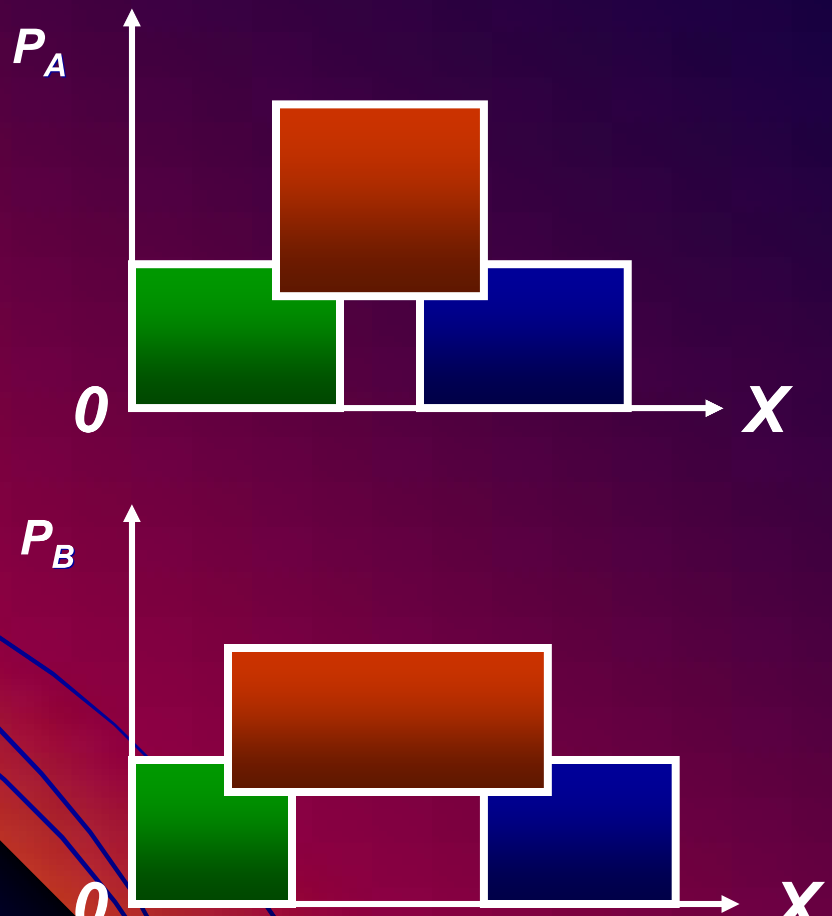
Extension Principle

$$\mu_{E(P^*)}(v) = \sup_g \left(\mu_{p_1}\left(\int_U \mu_{A_1}(u)g(u)du\right) \wedge \cdots \wedge \mu_{P_n}\left(\int_U \mu_{A_n}(u)g(u)du\right)\right)$$

subject to: $v = \int_U ug(u)du$

PERCEPTION-BASED DECISION ANALYSIS

ranking of f-granular probability distributions



maximization of expected utility ⟶ *ranking of fuzzy numbers*

USUALITY CONSTRAINT PROPAGATION RULE

X: random variable taking values in U
g: probability density of X

$$\frac{X \text{ isu } A}{\text{Prob }\{X \text{ is } B\} \text{ is } C}$$

X isu A ⟶ *Prob {X is A} is usually* ⟶

$$\pi(g) = \mu_{usually}\left(\int_U \mu_A(u)g(u)du\right)$$

$$\mu_C(v) = sup_g\left(\mu_{usually}\left(\int_U \mu_A(u)g(u)du\right)\right)$$

subject to:

$$v = \int_U \mu_B(u)g(u)du$$

CATEGORIES OF UNCERTAINTY

category 1: possibilistic

examples

crisp: $0 \leq X \leq a$;fuzzy: X is small

category 2: probabilistic

example

X isp N (m, σ^2)

category 3: possibility^2 (possibility of possibility) (type 2)

example:

grade of membership of μ in A is low

category 4: probabilistic^2 (probability of probability) (second order probability)

example: P(A) isp B

CONTINUED

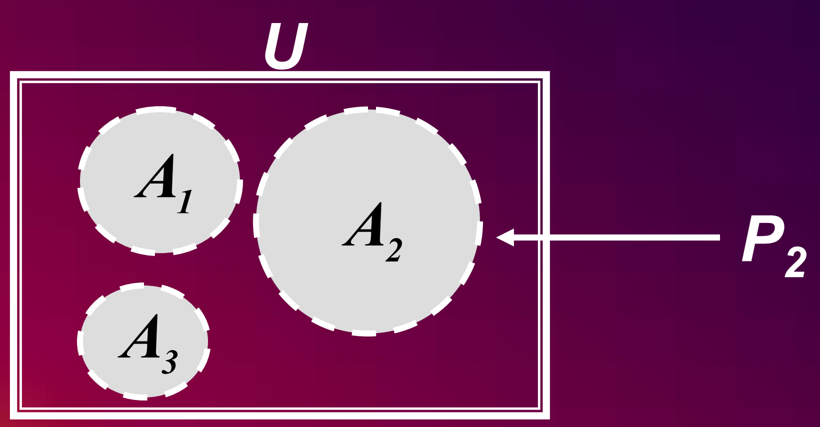
category 5: possibilistic\probabilistic (possibility of probability)

example:

X *isp* $(P_1\backslash A_1+\ldots+P_n\backslash A_n)$, Prob $\{X \text{ is } A_i\}$ is P_i

category 6: probabilistic\\possibilistic (probability of possibility)

X *isrs* $(P_1\backslash\backslash A_1+\ldots+P_n\backslash\backslash A_n)$

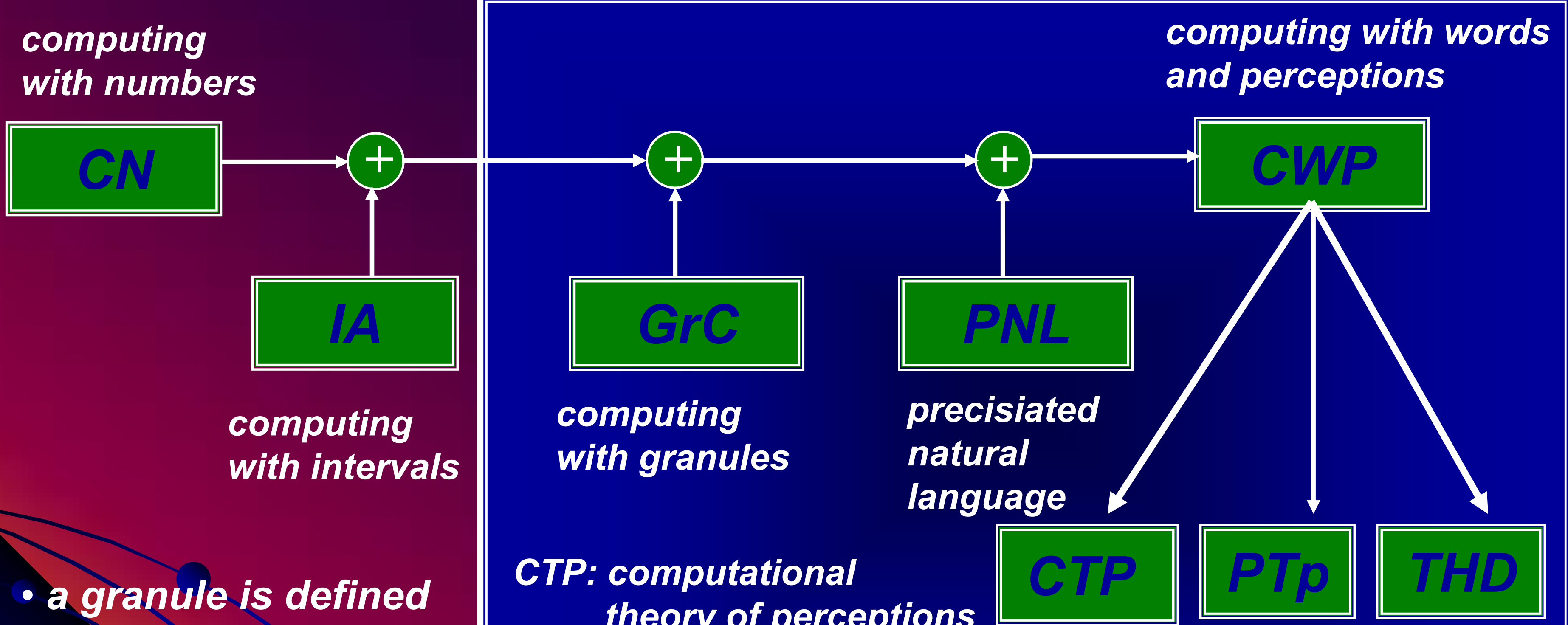


category 6 = fuzzy-set-valued granular probability distributions

New Tools

LAZ 4/24/2003

NEW TOOLS



- a granule is defined by a generalized constraint

CTP: computational theory of perceptions
PTp: perception-based probability theory
THD: theory of hierarchical definability

GRANULAR COMPUTING
GENERALIZED VALUATION
valuation = assignment of a value to a variable

X = 5	0 ≤ X ≤ 5	X is small	X isr R
point	interval	fuzzy interval	generalized

singular value

measurement-based

granular values

perception-based

PRECISIATED NATURAL LANGUAGE

PNL

CWP AND PNL

- *a concept which plays a central role in CWP is that of PNL (Precisiated Natural Language)*
- *basically, a natural language, NL, is a system for describing perceptions*
- *perceptions are intrinsically imprecise*
- *imprecision of natural languages is a reflection of the imprecision of perceptions*
- *the primary function of PNL is that of serving as a part of NL which admits precisiation*
- *PNL has a much higher expressive power than any language that is based on bivalent logic*

PRINCIPAL FUNCTIONS OF PNL

- *knowledge—and especially world knowledge—description language*
  - *Robert is tall*
  - *heavy smoking causes lung cancer*
- *definition language*
  - *smooth function*
  - *stability*
- *deduction language*

A is near B

B is near C

C is not far from A

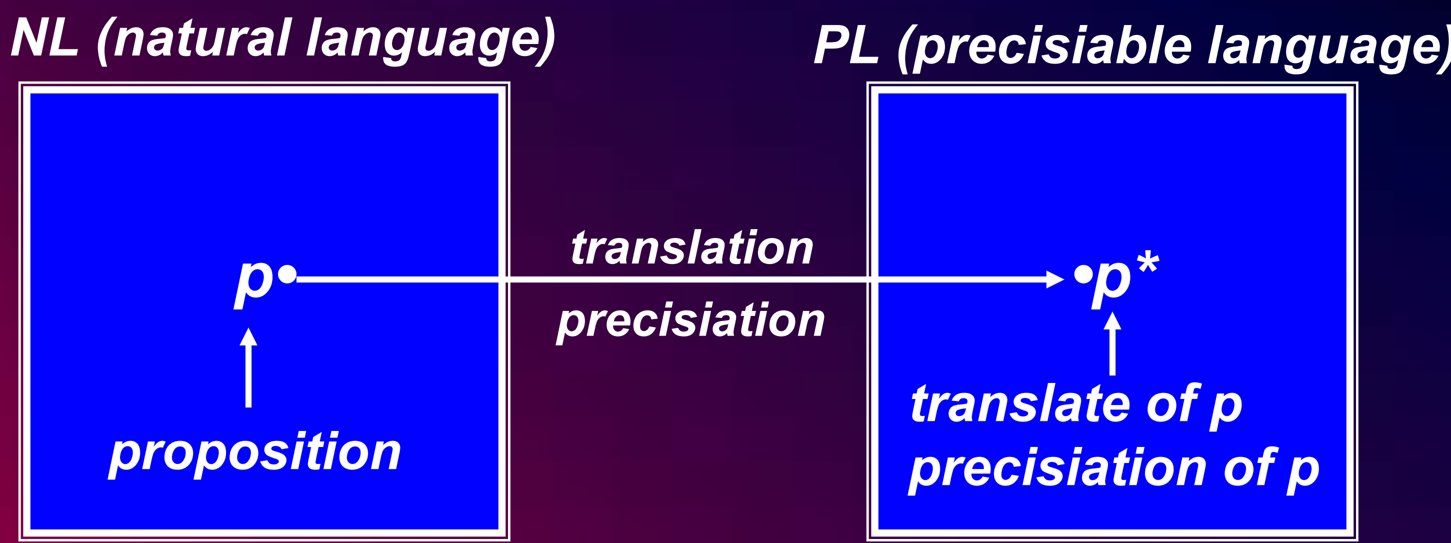
PNL

KEY POINTS

- *PNL is a subset of precisiable propositions/commands/questions in NL*
- *PNL is equipped with two dictionaries:*

  (1) from NL to GCL; and (2) from GCL to PFL; and (3) a modular multiagent deduction database (DDB) of rules of deduction (rules of generalized constrained propagation) expressed in PFL
- *the deduction database includes a collection of modules and submodules, among them the WORLD KNOWLEDGE module*

THE CONCEPT OF PRECISIATION



- p is precisiable w/r to PL = p is translatable into PL
- criterion of precisiability: p* is an object of computation

PL: propositional logic
predicate logic
modal logic
Prolog
LISP
SQL
•
Generalized Constraint Language (GCL) : p* = GC-form

PRECISIABILITY

- *Robert is tall: not PL-precisiable; PNL-precisiable*
- *all men are mortal: PL-precisiable*
- *most Swedes are tall: not PL-precisiable; PNL-precisiable*
- *about 20-25 minutes: not PL-precisiable; PNL-precisiable*
- *slow down: not PL-precisiable; PNL-precisiable*
- *overeating causes obesity: not PL-precisiable; PNL-precisiable*
- *Robert loves Anne: PNL-precisiable*
- *Robert loves women: not PNL-precisiable*
- *you are great: not PNL-precisiable*

PRECISIATION

- *precisiation is not coextensive with meaning representation*

  precisiation of p = precisiation of meaning of p

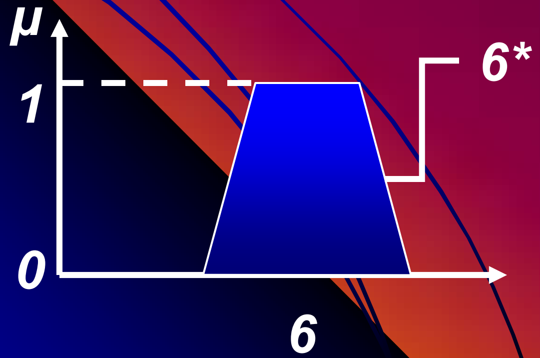
example:

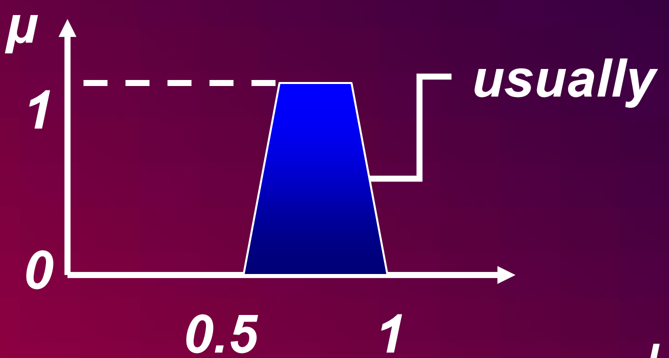
p = usually Robert returns from work at about 6pm.

I understand what you mean but can you be more precise?

yes

p $\longrightarrow$ Prob (Time (Robert.returns.from.work) is 6) is usually*





EXAMPLES

PL: propositional logic

- *Robert is taller than Alan* *taller (Robert, Alan)*

  Height (Robert)>Height (Alan)

PL: first-order predicate logic

- *all men are mortal* $\longrightarrow$ $\forall x$ *(man(x)* $\longrightarrow$ *mortal(x))*
- *most Swedes are tall* $\not\longrightarrow$ *not precisiable*

PL: PNL

most Swedes are tall $\longrightarrow$ Σ*Count (tall.Swedes/Swedes)*

is most

- *principal distinguishing features of PNL are:*

PL : GCL (Generalized Constraint Language)

DL (Deduction Logic): FL (fuzzy logic)

PNL is maximally expressive

THE CONCEPT OF A GENERALIZED CONSTRAINT (1985)

X isr R ← GC-form

- *R* — granular value of *X*; constraining relation
- *r* — modal variable (defines modality)
- *X* — constrained variable

principal modalities:

- *possibilistic (r = blank) : X is R , R=possibility distribution of X*
- *probabilistic (r = p) : X isp R : R=probability distribution of X*
- *veristic (r = v) : X isv R : R=verity (truth) distribution of X*
- *usuality (r=u) : X isu R : R=usual value of X*
- *random set (r=rs) : X isrs R : R=fuzzy-set-valued distribution of X*
- *fuzzy graph (r=fg) : X isfg : R=fuzzy graph of X*
- *bimodal (r=bm) : X isbm R : R=bimodal distribution of X*
- *Pawlak set (r=ps) : X isps R : upper and lower approximation to X*

GENERALIZED CONSTRAINT

- *standard constraint:* X ∈ C
- *generalized constraint:* *X isr R*

X isr R

copula

GC-form (generalized constraint form of type r)

type (modality) identifier

constraining relation

constrained variable

- *X =* $(X_1, \ldots, X_n)$
- *X may have a structure: X = Location (Residence(Carol))*
- *X may be a function of another variable: X = f(Y)*
- *X may be conditioned: (X/Y)*
- r := / ≤ / ... / ⊂ / ⊃ / blank / v / p / u / rs / fg / ps / ...

CONSTRAINT QUALIFICATION

- *constraint qualification: (X isr R) is q*

  qualifier

- *q*
  - *possibility*
  - *probability*
  - *verity (truth)*

- *example: (X is small) is unlikely*

INFORMATION: PRINCIPAL MODALITIES

- *possibilistic: r = blank*

  X is R (R: possibility distribution of X)

- *probabilistic: r = p*

  X isp R (R: probability distribution of X)

- *veristic: r = v*

  X isv R (R: verity (truth) distribution of X)

- *if r is not specified, default mode is possibilistic*

EXAMPLES (POSSIBILISTIC)

- *Eva is young* ⟶ *Age (Eva) is young*

  X R

- *Eva is much younger than Maria* ⟶
  (Age (Eva), Age (Maria)) is much younger

  X R

- *most Swedes are tall*
  ⟶ *ΣCount (tall.Swedes/Swedes) is most*

  X R

EXAMPLES (PROBABILISITIC)

- *X is a normally distributed random variable with mean m and variance σ^2* ⟶

  $$X \text{ isp } N(m, \sigma^2)$$

- *X is a random variable taking the values u_1, u_2, u_3 with probabilities p_1, p_2 and p_3, respectively* ⟶

  $$X \text{ isp } (p_1\backslash u_1 + p_2\backslash u_2 + p_3\backslash u_3)$$

EXAMPLES (VERISTIC)

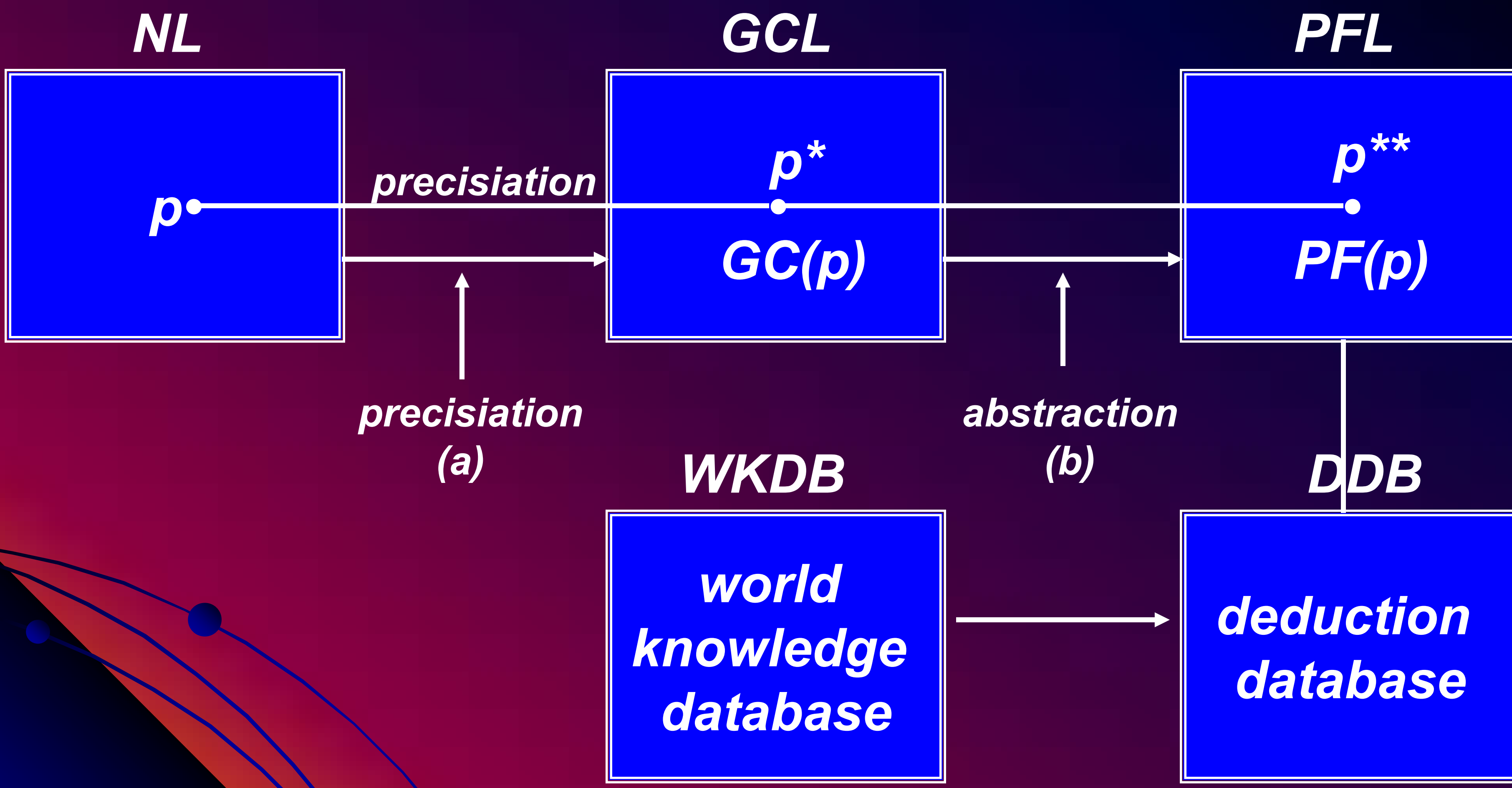
- *Robert is half German, quarter French and quarter Italian*

  Ethnicity (Robert) isv (0.5|German + 0.25|French + 0.25|Italian)

- *Robert resided in London from 1985 to 1990*

  Reside (Robert, London) isv [1985, 1990]

BASIC STRUCTURE OF PNL



•In PNL, deduction=generalized constraint propagation
DDB: deduction database=collection of protoformal rules governing generalized constraint propagation
WKDB: PNL-based

EXAMPLE OF TRANSLATION

- *P: usually Robert returns from work at about 6 pm*
- *P*: Prob {(Time(Return(Robert)) is 6 pm} is usually*
- *PF(p): Prob {X is A} is B*
- *X: Time (Return (Robert))*
- *A: 6 pm*
- *B: usually*

$p \in NL$

$p^* \in GCL$

$PF(p) \in PFL$

BASIC STRUCTURE OF PNL

DICTIONARY 1

NL	GCL
p	GC(p)

DICTIONARY 2

GCL	PFL
GC(p)	PF(p)

MODULAR DEDUCTION DATABASE

POSSIBILITY MODULE

PROBABILITY MODULE

agent

FUZZY ARITHMETIC MODULE

RANDOM SET MODULE

FUZZY LOGIC MODULE

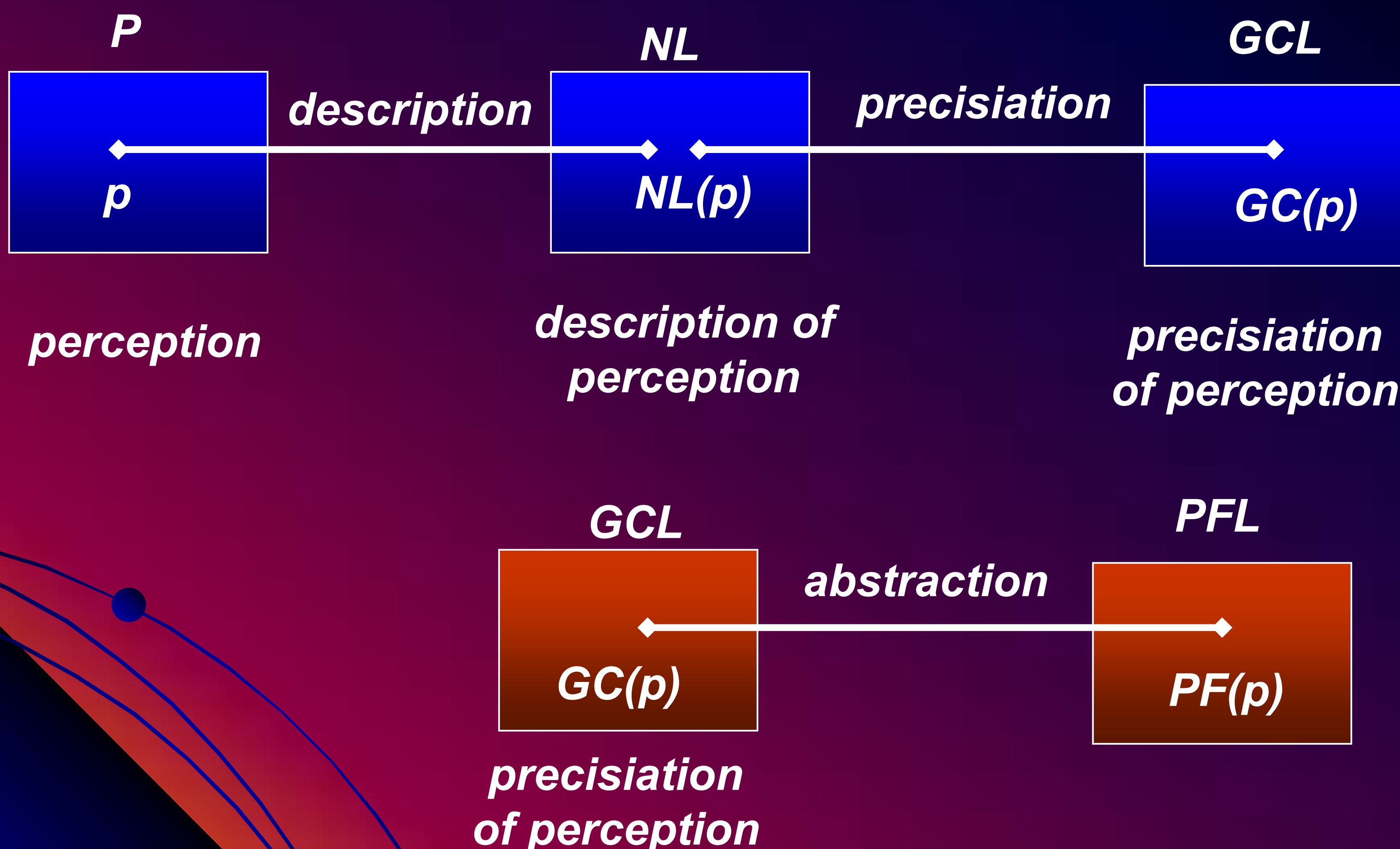
EXTENSION PRINCIPLE MODULE

GENERALIZED CONSTRAINT LANGUAGE (GCL)

- GCL is generated by combination, qualification and propagation of generalized constraints
- in GCL, rules of deduction are the rules governing generalized constraint propagation
- examples of elements of GCL
  - (X isp R) and (X,Y) is S)
  - (X isr R) is unlikely) and (X iss S) is likely
  - if X is small then Y is large

- the language of fuzzy if-then rules is a sublanguage of PNL

THE BASIC IDEA



GCL (Generalized Constrain Language) is maximally expressive

DICTIONARIES

1:

proposition in NL	precisiation
p	p* (GC-form)
most Swedes are tall	Σ Count (tall.Swedes/Swedes) is most

2:

precisiation	protoform
p* (GC-form)	PF(p*)
Σ Count (tall.Swedes/Swedes) is most	Q A's are B's

TRANSLATION FROM NL TO PFL

examples

Eva is young ⟶ *A (B) is C*

Age *Eva* *young*

Eva is much younger than Pat ⟶ *(A (B), A (C)) is R*

Age *Eva* *Age* *Pat* *much younger*

usually Robert returns from work at about 6pm ⟶

Prob {A is B} is C

usually

about 6 pm

Time (Robert returns from work)

PNL AS A

DEFINITION LANGUAGE

LAZ 4/24/2003

HIERARCHY OF DEFINITION LANGUAGES

NL: natural language

B language: standard mathematical bivalent-logic-based language

F language: fuzzy logic language without granulation

F.G language: fuzzy logic language with granulation

PNL: Precisiated Natural Language

Note: the language of fuzzy if-then rules is a sublanguage of PNL

Note: a language in the hierarchy subsumes all lower languages

SIMPLIFIED HIERARCHY

PNL	← fuzzy-logic-based
B language	← bivalent-logic-based
NL	

The expressive power of the B language – the standard bivalence-logic-based definition language – is insufficient

Insufficiency of the expressive power of the B language is rooted in the fundamental conflict between bivalence and reality

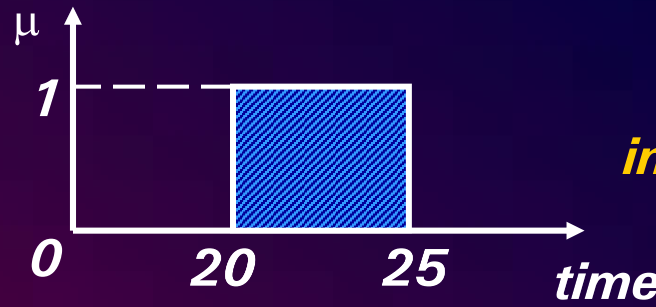
EVERYDAY CONCEPTS WHICH CANNOT BE DEFINED REALISTICALY THROUGH THE USE OF B

- *check-out time is 12:30 pm*
- *speed limit is 65 mph*
- *it is cloudy*
- *Eva has long hair*
- *economy is in recession*
- *I am risk averse*
- …

PRECISIATION/DEFINITION OF PERCEPTIONS

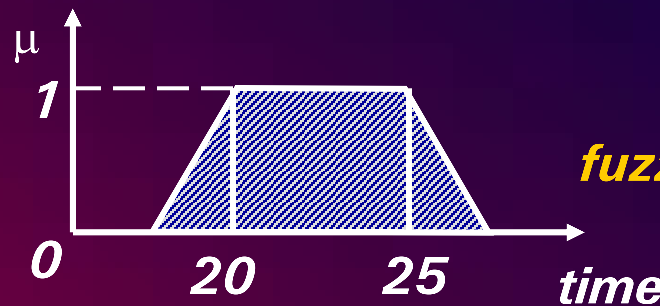
Perception: ABOUT 20-25 MINUTES

B definition:



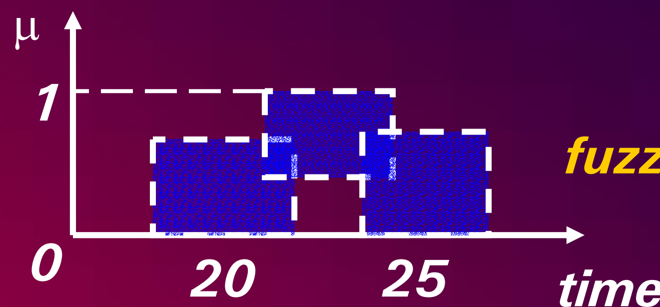
interval

F definition:



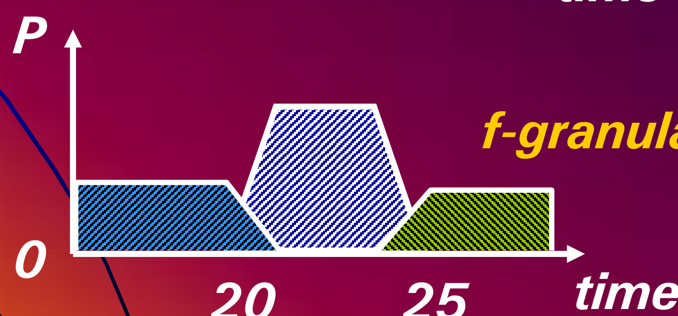
fuzzy interval

F.G definition:



fuzzy graph

PNL definition:



f-granular probability distribution

INSUFFICIENCY OF THE B LANGUAGE

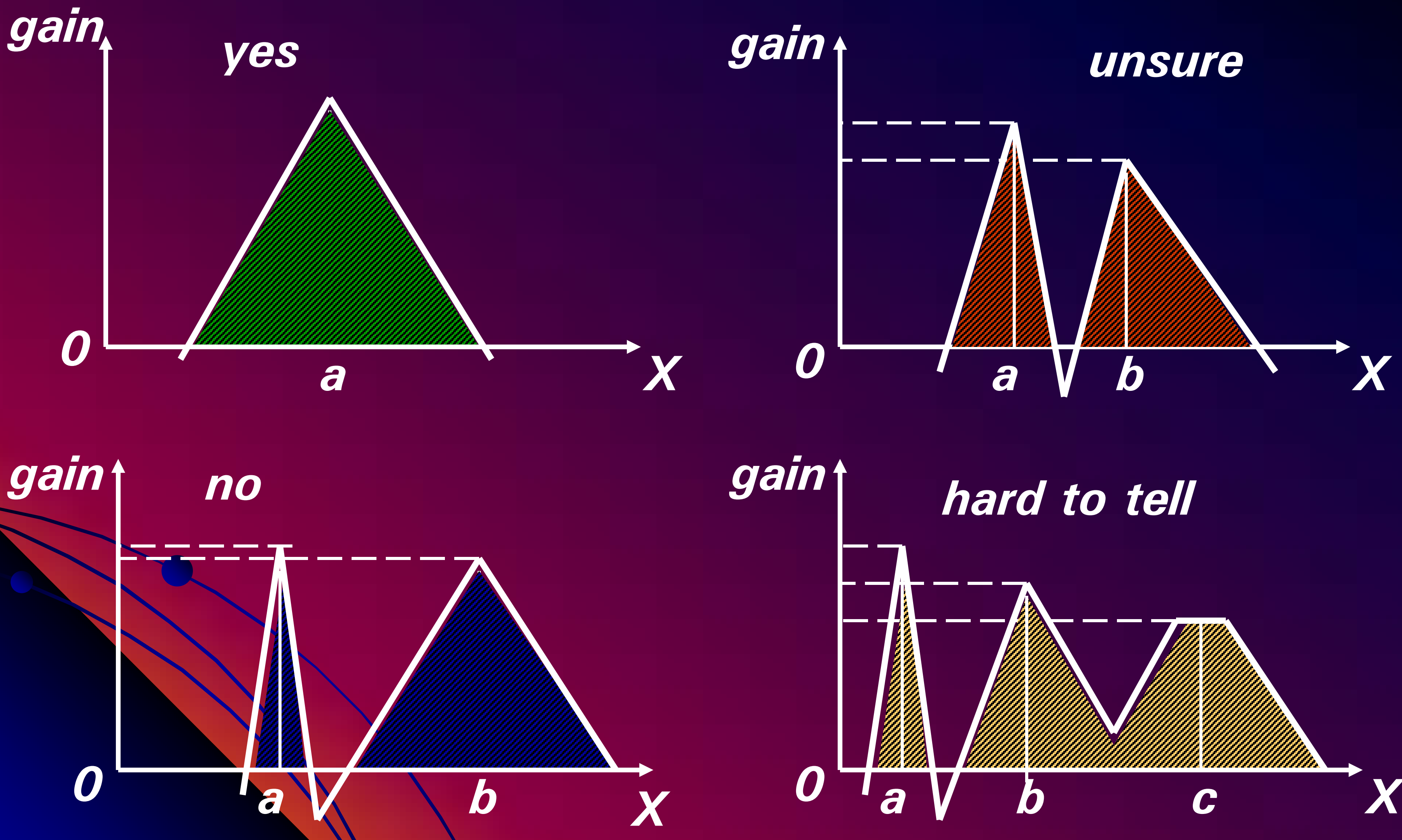
Concepts which cannot be defined

- ***causality***
- ***relevance***
- ***intelligence***

Concepts whose definitions are problematic

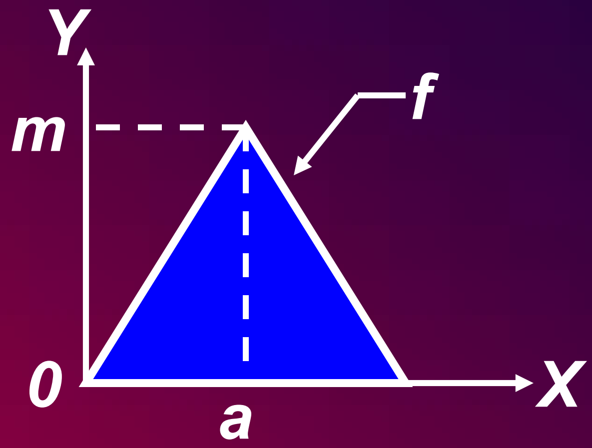
- ***stability***
- ***optimality***
- ***statistical independence***
- ***stationarity***

DEFINITION OF OPTIMALITY
OPTIMIZATION = MAXIMIZATION?



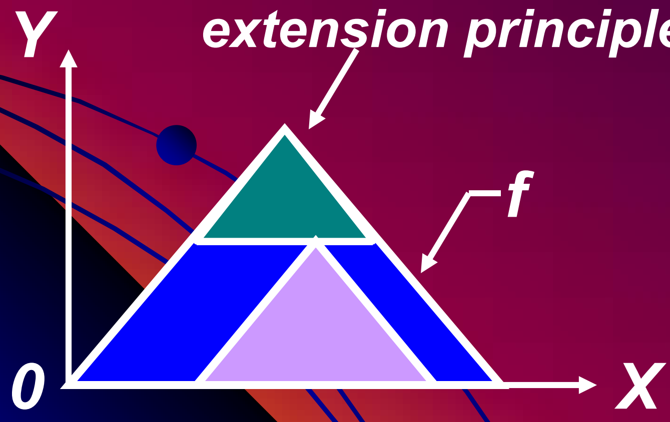
- *definition of optimal X requires use of PNL*

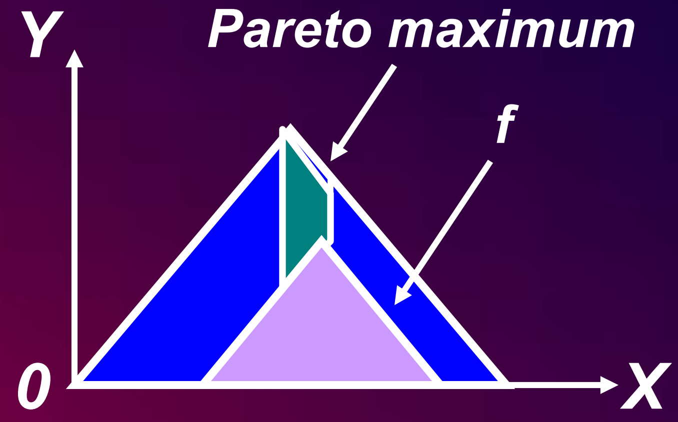
MAXIMUM ?



a) $\forall x\ (f(x) \leq f(a))$

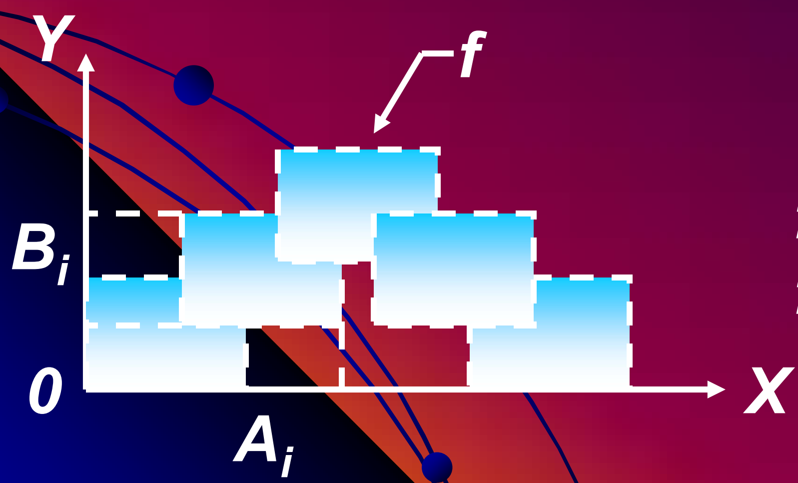
b) $\sim (\exists x\ (f(x) > f(a))$





b) $\sim (\exists x\ (f(x)$ dominates $f(a))$

MAXIMUM ?



$f = \Sigma_i A_i \times B_i$

f: if X is A_i then Y is B_i, $i=1, \ldots, n$

EXAMPLE

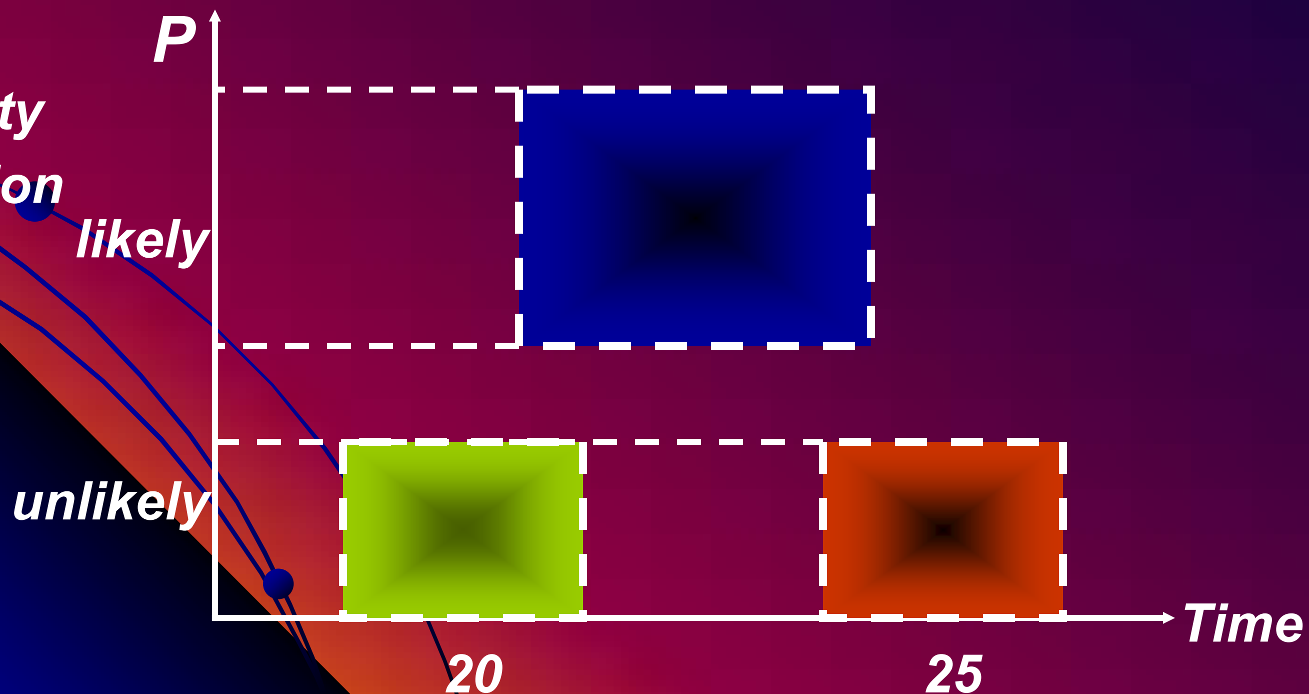
- *I am driving to the airport. How long will it take me to get there?*
- *Hotel clerk's perception-based answer: about 20-25 minutes*
- *"about 20-25 minutes" cannot be defined in the language of bivalent logic and probability theory*
- *To define "about 20-25 minutes" what is needed is PNL*

EXAMPLE

PNL definition of “about 20 to 25 minutes”

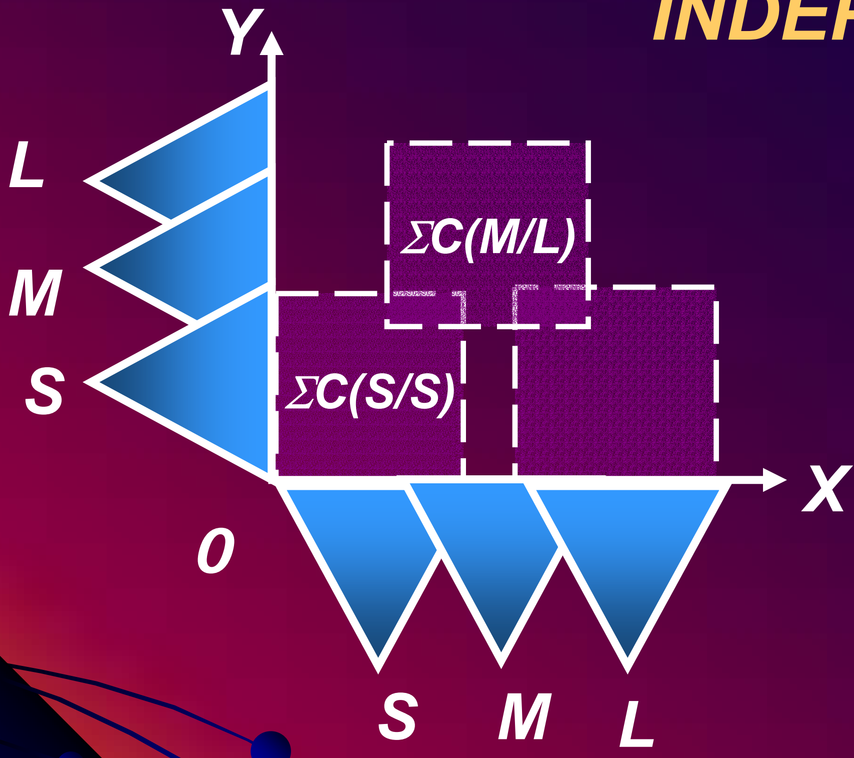
Prob {getting to the airport in less than about 25 min} is unlikely
Prob {getting to the airport in about 20 to 25 min} is likely
Prob {getting to the airport in more than 25 min} is unlikely

granular probability distribution



LAZ 4/24/2003

PNL-BASED DEFINITION OF STATISTICAL INDEPENDENCE



contingency table

3	L/S	L/M	L/L
2	M/S	M/M	M/L
1	S/S	S/M	S/L
	1	**2**	**3**

$$\Sigma\,(M/L) = \frac{\Sigma C\,(M \times L)}{\Sigma C\,(L)}$$

- *degree of independence of Y from X=*
  degree to which columns 1, 2, 3 are identical

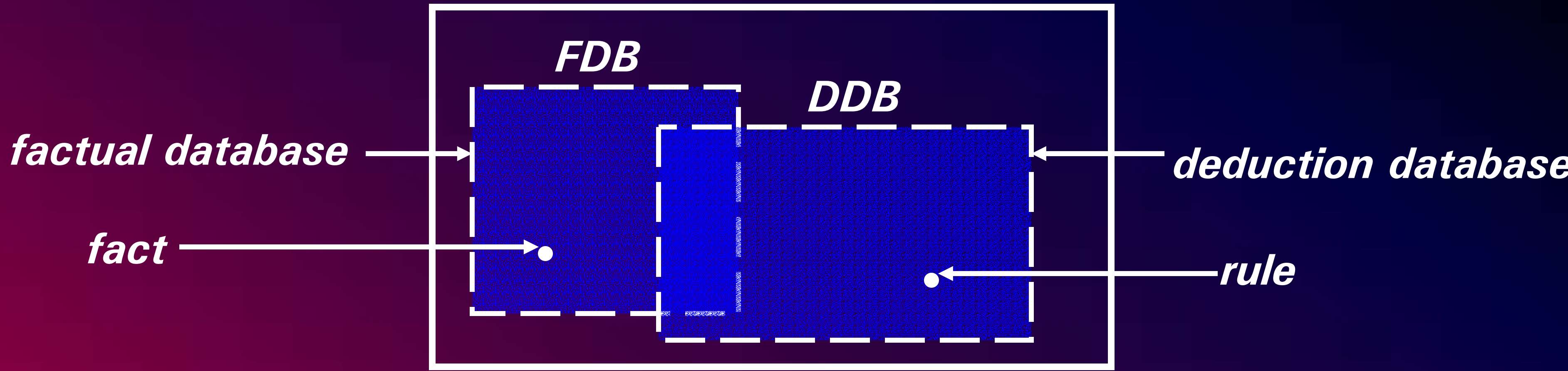
→ *PNL-based definition*

PROTOFORM LANGUAGE

PFL

LAZ 4/24/2003

ORGANIZATION OF KNOWLEDGE



knowledge — measurement-based / perception-based

- much of human knowledge is perception-based

examples of factual knowledge

- height of Eiffel Tower is 324 m (with antenna) (measurement-based)
- Berkeley is near San Francisco (perception-based)
- icy roads are slippery (perception-based)
- if Marina is a student then it is likely that Marina is young (perception-based)

THE CONCEPT OF PROTOFORM

- *a protoform is an abstracted prototype of a class of propositions*

examples:

most Swedes are tall $\xrightarrow{\text{P-abstraction}}$ *Q A's are B's*

many Americans are foreign-born

overeating causes obesity $\xrightarrow{\text{P-abstraction}}$ *Q A's are B's*

obesity is caused by overeating $\xrightarrow{\text{P-abstraction}}$ *Q B's are A's*

THE CONCEPT OF PROTOFORM

KEY POINTS

- *protoform: abbreviation of "prototypical form"*
- *PF(p): protoform of p*
- *PF(p): deep semantic structure of p*
- *PF(p): abstraction of precisiation of p*
- *abstraction is a form of summarization*
- *if p has a logical form, LF(p), then PF(p) is an abstraction of LF(p)*

all men are mortal $\longrightarrow \forall x(\text{man}(x) \longrightarrow \text{mortal}(x)) \longrightarrow \forall x(A(x) \longrightarrow B(x))$

$\forall x(\text{man}(x) \longrightarrow \text{mortal}(x))$: LF

$\forall x(A(x) \longrightarrow B(x))$: PF

CONTINUED

- *if p does not have a logical form but has a generalized constraint form, GC(p), then PF(p) is an abstraction of GC(p)*

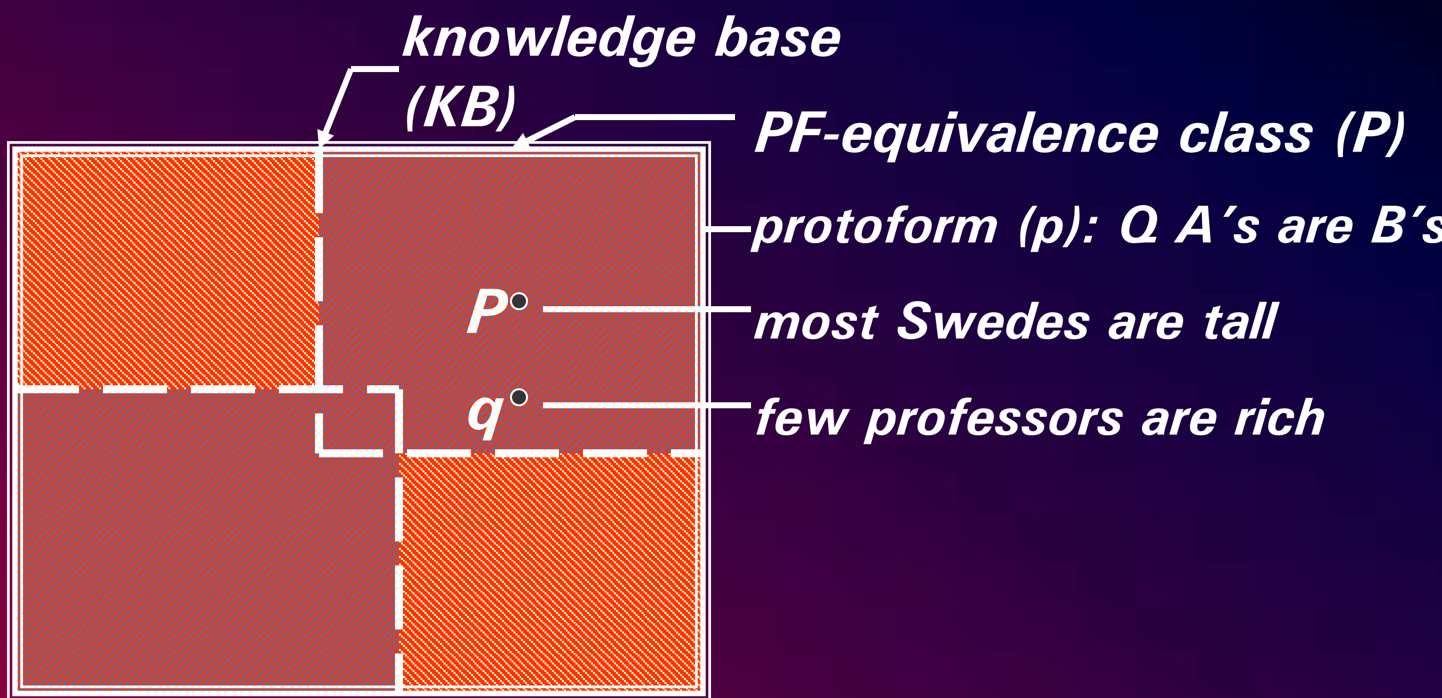
most Swedes are tall $\longrightarrow \Sigma Count(tall.Swedes / Swedes)$ *is most*

GC(p)

QA's are B's

PF(p)

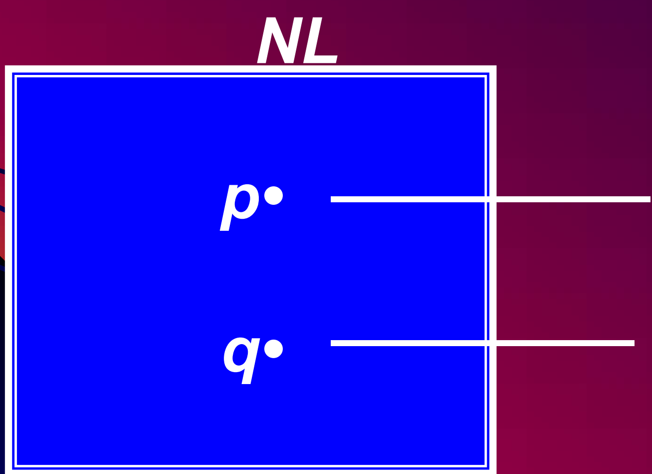
PROTOFORM AND PF-EQUIVALENCE



- P is the class of PF-equivalent propositions
- P does not have a prototype
- P has an abstracted prototype: Q A′s are B′s
- P is the set of all propositions whose protoform is: Q A′s are B′s

CONTINUED

- ***abstraction has levels, just as summarization does***
- ***p and q are PF-equivalent at level α if at level of abstraction α, PF(p)=PF(q)***



most Swedes are tall → *QA's are B's*

a few professors are rich → *QA's are B's*

DEDUCTION (COMPUTING) WITH PERCEPTIONS

deduction

$$\begin{array}{c} p_1 \\ p_2 \\ \bullet \\ \bullet \\ p_n \\ \hline p_{n+1} \end{array}$$

example

Dana is young
Tandy is a few years older than Dana
Tandy is (young + few)

deduction with perceptions involves the use of protoformal rules of generalized constraint propagation

DEDUCTION MODULE

- *rules of deduction are rules governing generalized constraint propagation*
- *rules of deduction are protoformal*

examples

generalized modus ponens

X is A
if X is B then Y is C

Y is A $\circ$ *(B* $\longrightarrow$ *C)*

$$\mu_y(v) = \sup(\mu_A(u) \wedge \mu_{B\rightarrow C}(u,v))$$

Prob (A) is B

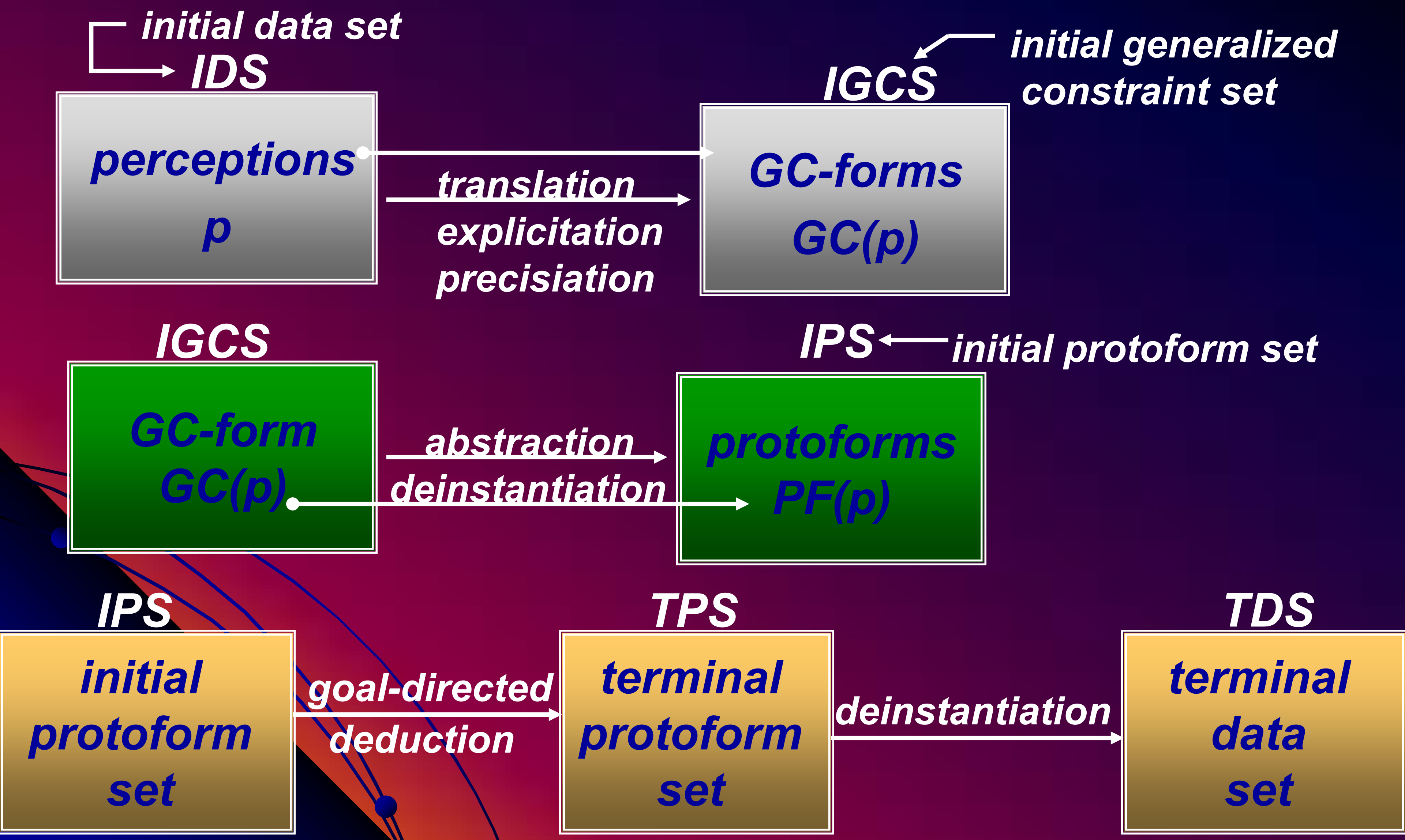
Prob (C) is D

$$\mu_D(v) = \sup_g(\mu_B(\int_U \mu_A(u)g(u)du))$$

subject to

$$v = \int_U \mu_C(u)g(u)du$$

REASONING WITH PERCEPTIONS: DEDUCTION MODULE



LAZ 4/24/2003

PROTOFORMAL CONSTRAINT PROPAGATION

p	*GC(p)*	*PF(p)*
Dana is young	*Age (Dana) is young*	*X is A*
Tandy is a few years older than Dana	*Age (Tandy) is (Age (Dana)) + few*	*Y is (X + B)*

$$\frac{\begin{array}{c} X \text{ is } A \\ Y \text{ is } (X+B) \end{array}}{Y \text{ is } A+B} \longrightarrow \boxed{\text{Age (Tandy) is (young + few)}}$$

$$\mu_{A+B}(v) = \sup_u(\mu_A(u) + \mu_B(v - u))$$

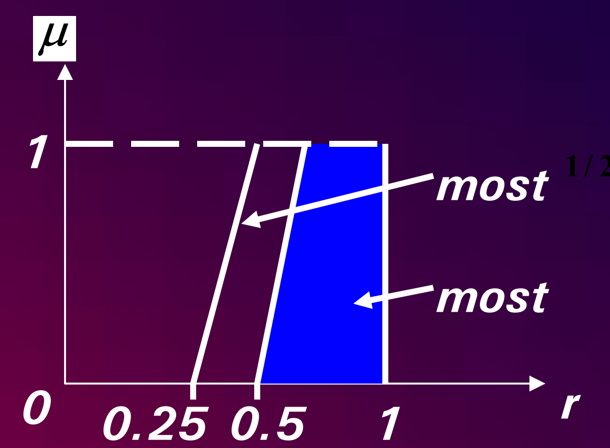
EXAMPLE OF DEDUCTION

$$\frac{\textit{most Swedes are tall}}{\textit{? R Swedes are very tall}}$$

most Swedes are tall $\xrightarrow{\textit{s/a-transformation}}$ *Q A′s are B′s*

$$\frac{\textit{Q A′s are B′s}}{Q^{1/2}\ \textit{A′s are}\ ^2\textit{B′s}}$$

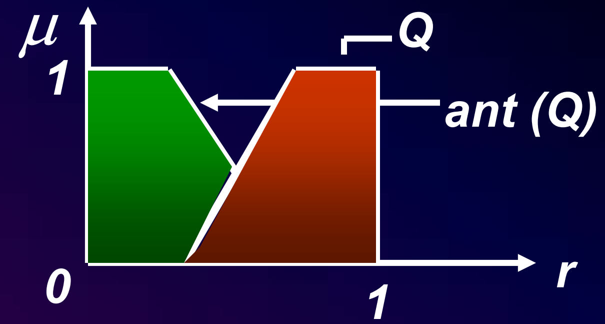
most$^{1/2}$ *Swedes are very tall*



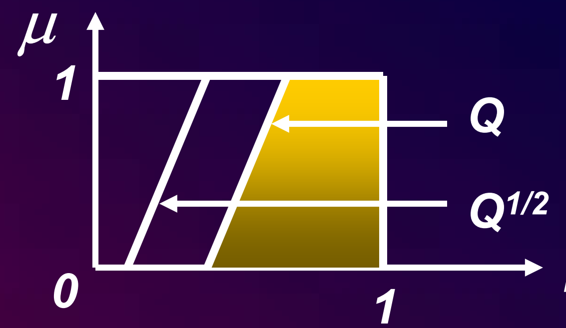
COUNT-AND MEASURE-RELATED RULES

crisp

$$\frac{Q\ A\text{'s are } B\text{'s}}{ant\ (Q)\ A\text{'s are not } B\text{'s}}$$



$$\frac{Q\ A\text{'s are } B\text{'s}}{Q^{1/2}\ A\text{'s are } {}^2B\text{'s}}$$



$$\frac{\text{most Swedes are tall}}{\text{ave (height) Swedes is ?h}} \longrightarrow \frac{Q\ A\text{'s are } B\text{'s}}{\text{ave } (B|A) \text{ is ?C}}$$

$$\mu_{ave}(v) = sup_a\, \mu_Q\left(\frac{1}{N}\sum_i \mu_B(a_i)\right) \quad , a = (a_1, \ldots, a_N)$$

$$v = \frac{1}{N}\left(\sum_i a_i\right)$$

CONTINUED

$$\text{not}(QA\text{'s are }B\text{'s}) \longleftrightarrow (\text{not }Q)\ A\text{'s are }B\text{'s}$$

$$\frac{\begin{array}{l} Q_1\ \ A\text{'s are }B\text{'s} \\ Q_2\ \ (A\&B)\text{'s are }C\text{'s} \end{array}}{Q_1\ \ Q_2\ \ A\text{'s are }(B\&C)\text{'s}}$$

$$\frac{\begin{array}{l} Q_1\ \ A\text{'s are }B\text{'s} \\ Q_2\ \ A\text{'s are }C\text{'s} \end{array}}{(Q_1 + Q_2\ -1)\ A\text{'s are }(B\&C)\text{'s}}$$

INTERPOLATION

$$\pi(g) = \mu_{Pi(1)}\left(\int_U \mu_{A_i}(u)g(u)du\right) \wedge \cdots \wedge \mu_{Pi(n)}\left(\int_U \mu_{A_n}(u)g(u)du\right)$$

$$\pi^*\left(\int_U \mu_A(u)g(u)du\right) \quad \text{is ?A}$$

$$\pi^*(v) = \sup_g \mu_{Pi(1)}\left(\int_U \mu_{A_i}(u)g(u)du\right) \wedge \cdots \wedge \mu_{Pi(n)}\left(\int_U \mu_{A_n}(u)g(u)du\right)$$

subject to: $v = \int_U \mu_A(u)g(u)du$

$$\int_U g(u)du = 1$$

CONTINUED

$\Pi(g)$: possibility distribution of g

$\Pi(g)$: $\mu_{Pi(1)}(\int_U \mu_{A_i}(u)g(u)du) \wedge \cdots \wedge \mu_{Pi(n)}(\int_U \mu_{A_n}(u)g(u)du)$

extension principle

$$\frac{\Pi(g)}{\Pi^*(f(g))}$$

$\Pi^*(v) = sup_g(\Pi(g))$

subject to:

$v = f(g)$

EXPECTED VALUE

$$\pi(g) = \mu_{Pi(1)}\left(\int_U \mu_{A_i}(u)g(u)du\right) \wedge \cdots \wedge \mu_{Pi(n)}\left(\int_U \mu_{A_n}(u)g(u)du\right)$$

$\pi^*\left(\int_U ug(u)du\right)$ is ?A

$$\pi^*(v) = \sup_g \mu_{Pi(1)}\left(\int_U \mu_{A_i}(u)g(u)du\right) \wedge \cdots \wedge \mu_{Pi(n)}\left(\int_U \mu_{A_n}(u)g(u)du\right)$$

subject to: $v = \int_U ug(u)du$

CONTINUED

- Prob {X is A_i} is $P_{j(i)}$, $i=1, \ldots, m$, $j=1, \ldots, n$
- $\int g(u)du=1$
- G is small $\longrightarrow$ $\forall u(g(u)$ is small)

Prob {X is A} is ?v

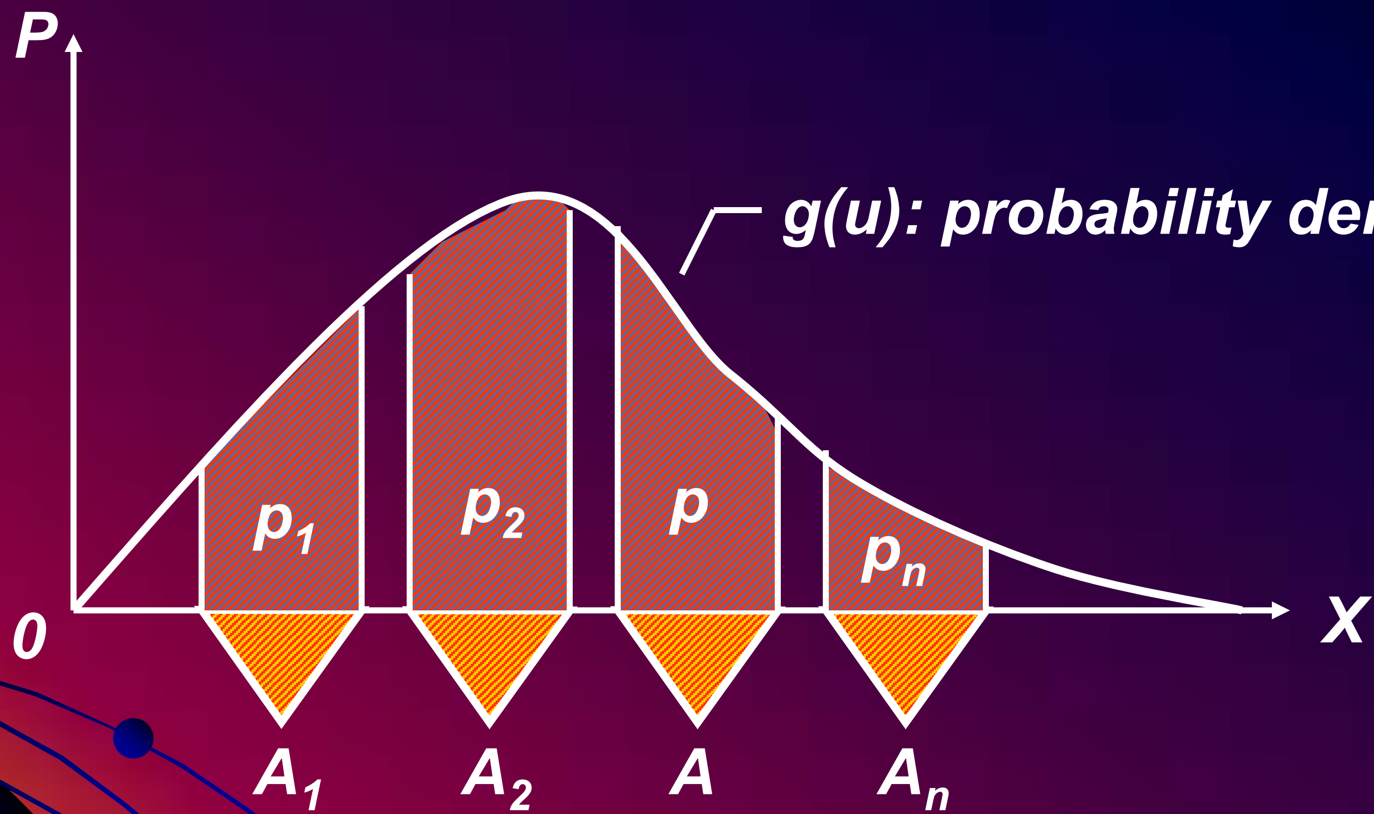
Prob {X is A_i} = $\int_U g(u)\mu_{A_i}(u)du$

construct: $\mu_{P_{j(i)}}(v) = \mu_{P_{j(i)}}(\int_U g(u)\mu_{A_i}(u)du)$

PROBABILITY MODULE

LAZ 4/24/2003

INTERPOLATION OF BIMODAL DISTRIBUTIONS



p_i is P_i : granular value of p_i , i=1, …, n

(P_i , A_i) , i=1, …, n are given

A is given

(?P, A)

INTERPOLATION MODULE AND PROBABILITY MODULE

Prob {X is A_i} is P_i , i = 1, …, n

Prob {X is A} is Q

$$\mu_Q(v) = sup_g(\mu_{P_1}(\int_U \mu_{A_1}(u)g(u)du) \wedge \cdots \wedge$$

$$\mu_{P_n} \int_U \mu_{P_n}(\int_U \mu_{A_n}(u)g(u)du))$$

subject to

$$U = \int_U \mu_A(u)g(u)du$$

PROBABILISTIC CONSTRAINT PROPAGATION RULE
(a special version of the generalized extension principle)

$$\int_U g(u)\mu_A(u)du \quad \textit{is } R$$

$$\int_U g(u)\mu_B(u)du \quad \textit{is } ?S$$

$$\mu_S(v) = \sup_g(\mu_R(\int_U g(u)\mu_A(u)du))$$

subject to

$$v = \int_U g(u)\mu_B(u)du$$

$$\int_U g(u)du = 1$$

USUALITY SUBMODULE

CONJUNCTION

$$\frac{X \text{ is } A \quad X \text{ is } B}{X \text{ is } A \cap B}$$

$$\frac{X \text{ isu } A \quad X \text{ isu } B}{X \text{ isr } A \cap B}$$

- determination of r involves interpolation of a bimodal distribution

USUALITY — QUALIFIED RULES

$$\frac{X \text{ isu } A}{X \text{ isun (not } A)}$$

$$\frac{\begin{array}{c} X \text{ isu } A \\ Y = f(X) \end{array}}{Y \text{ isu } f(A)}$$

$$\mu_{f(A)}(v) = \sup_{u|v=f(u)}(\mu_A(u))$$

USUALITY — QUALIFIED RULES

X isu A

Y isu B

Z = f(X,Y)

Z isu f(A, B)

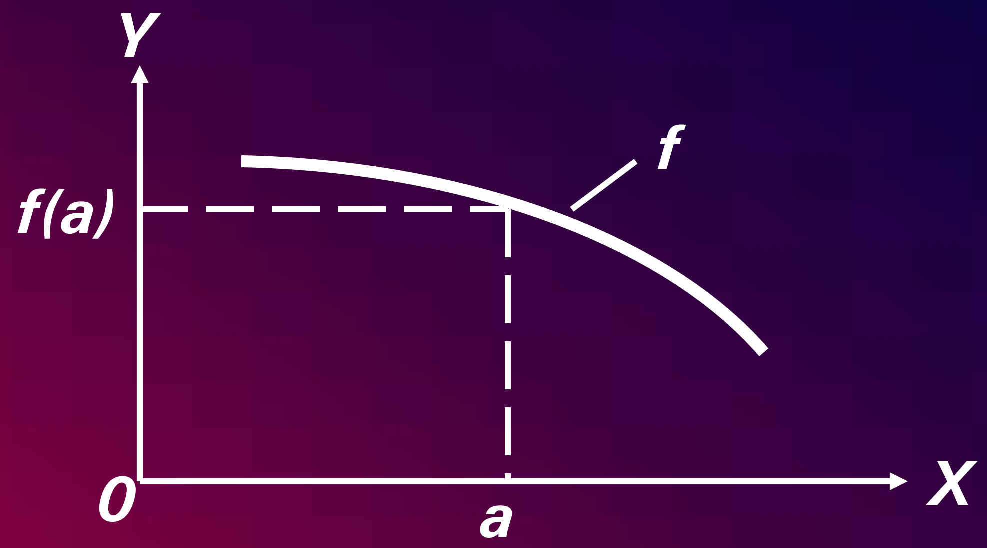
$$\mu_Z(w) = \sup_{u,v|w=f(u,v)}(\mu_A(u) \wedge \mu_B(v))$$

EXTENSION PRINCIPLE MODULE

LAZ 4/24/2003

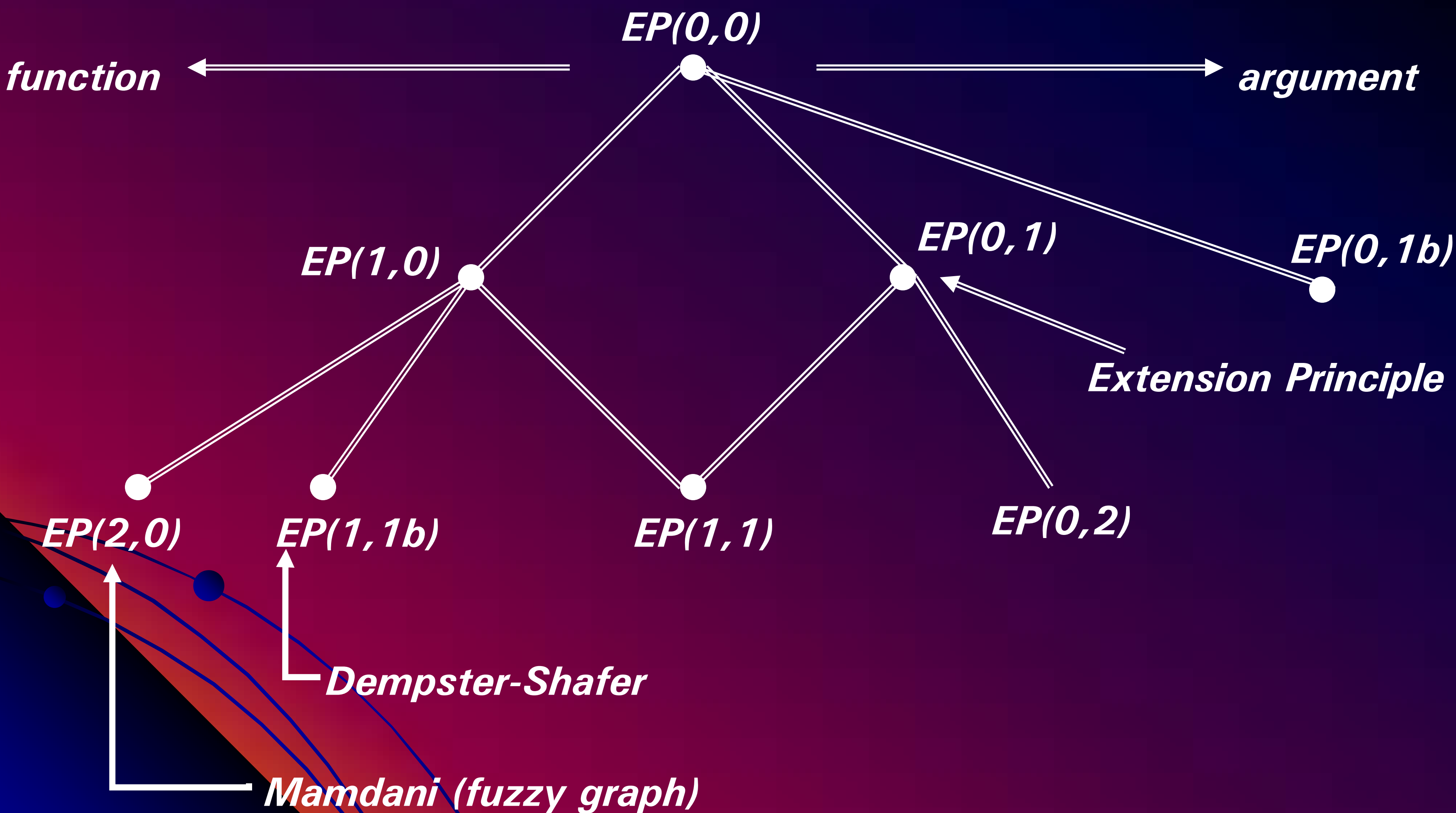
PRINCIPAL COMPUTATIONAL RULE IS THE EXTENSION PRINCIPLE (EP)

point of departure: function evaluation

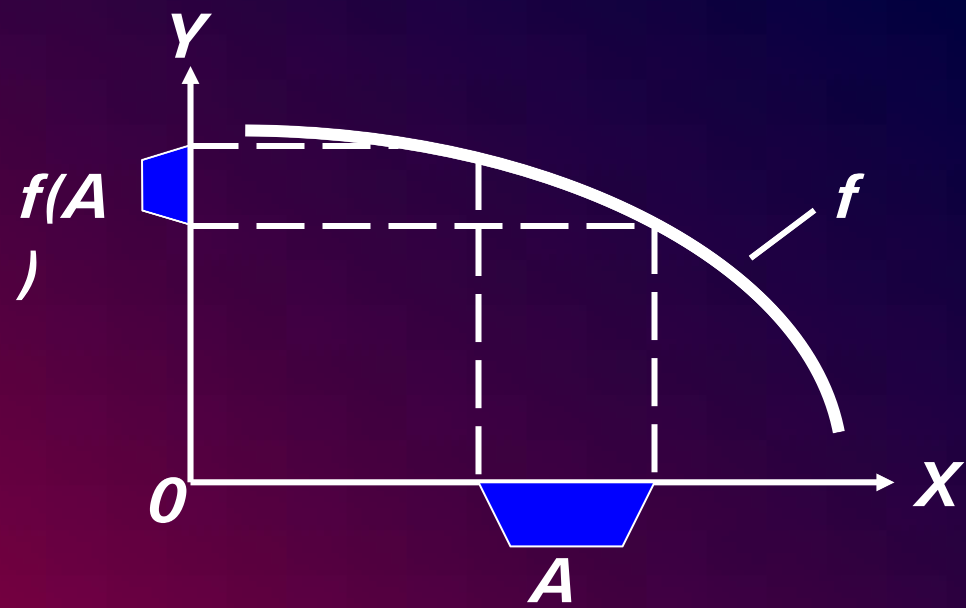


$$\frac{\begin{gathered}X = a\\ Y = f(X)\end{gathered}}{Y = f(a)}$$

EXTENSION PRINCIPLE HIERARCHY



VERSION EP(0,1) (1965; 1975)



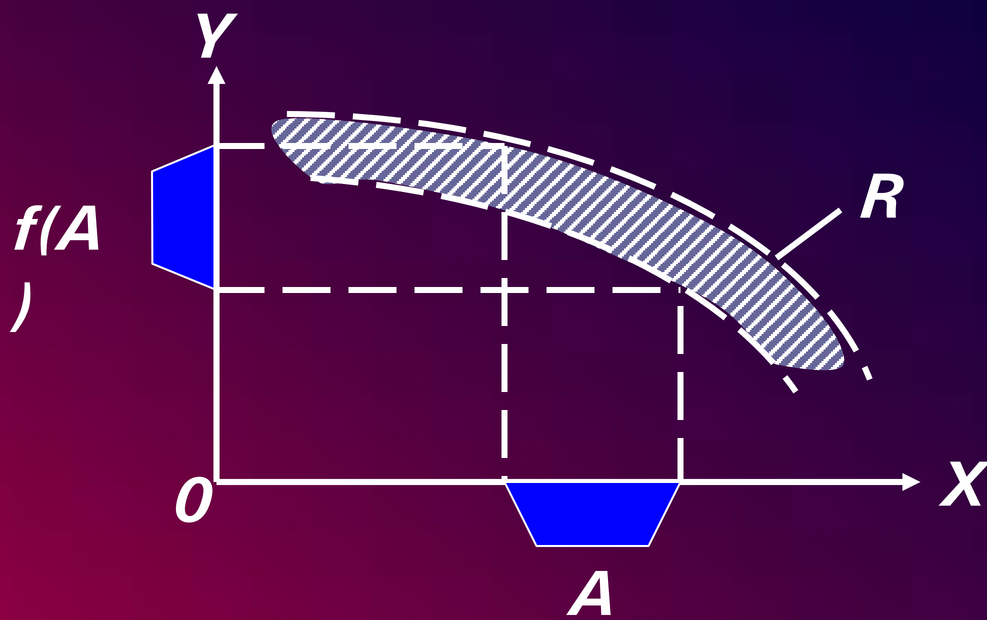
$$\frac{X \text{ is } A \quad Y = f(X)}{Y = f(A)}$$

$$\mu_{f(A)}(v) = \sup_u (\mu_A(u))$$

subject to

$$v = f(u)$$

VERSION EP(1,1) (COMPOSITIONAL RULE OF INFERENCE) (1965)

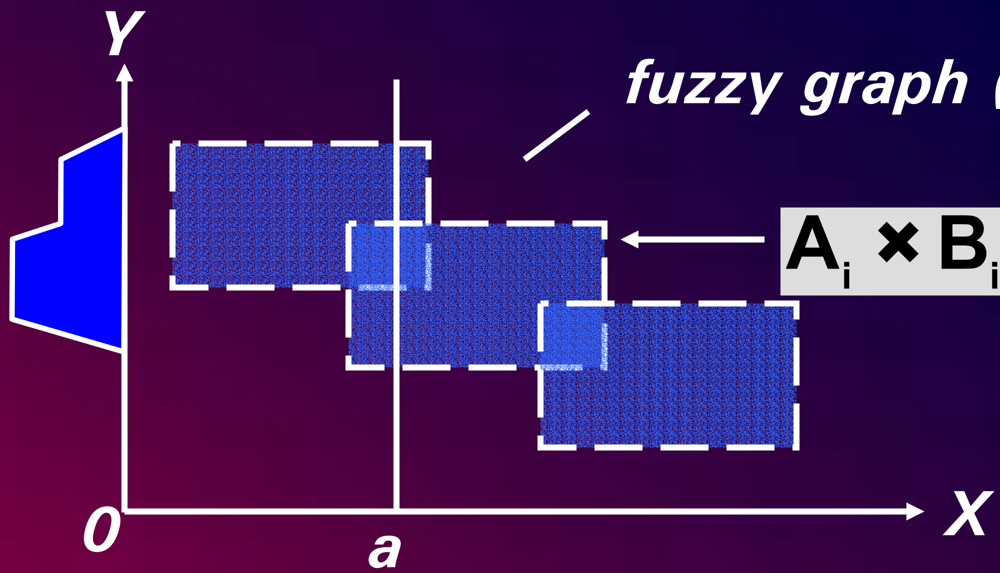


X is A
(X,Y) is R

Y is A ∘ R

$$\mu_Y(v) = \sup_u(\mu_A(u) \wedge \mu_R(u,v))$$

EXTENSION PRINCIPLE EP(2,0) (Mamdani)



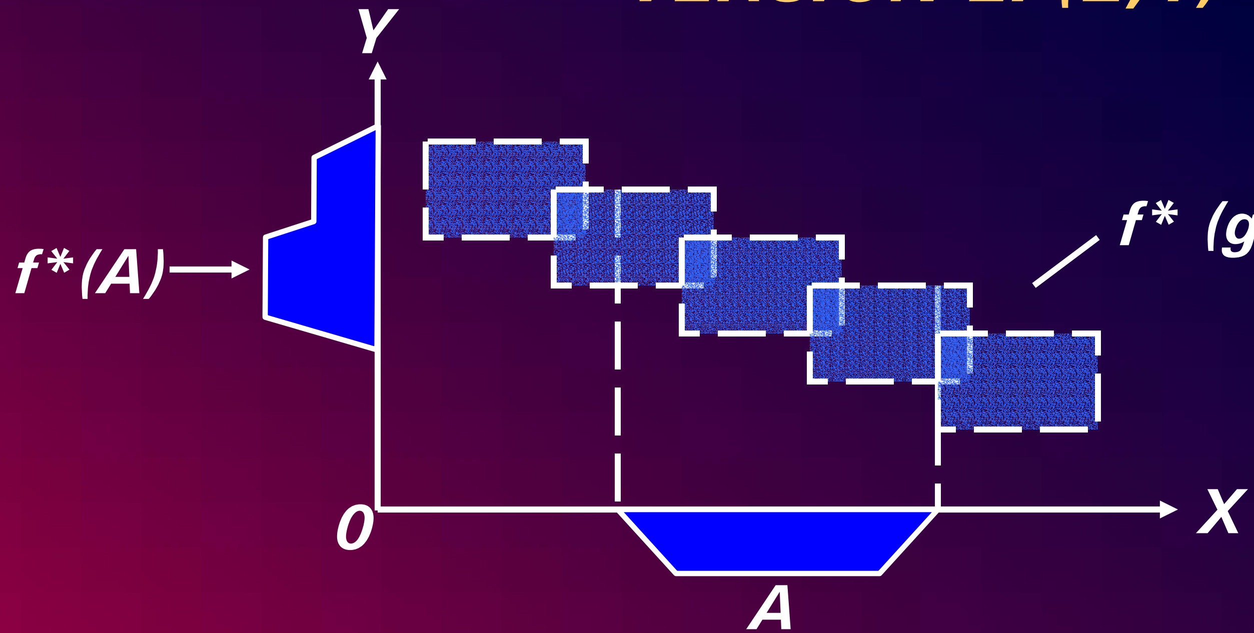
$f^* = \sum_i A_i \times B_i$

$X = a$

$Y = \sum_i \mu_{A_i}(a) \wedge B_i$

(if X is A_i then Y is B_i)

VERSION EP(2,1)



$$\frac{\begin{array}{c} X \text{ is } A \\ (X, Y) \text{ is } R \end{array}}{Y \text{ is } \Sigma_i m_i \wedge B_i}$$

$$R = \Sigma_i A_i \times B_i$$

$m_i = sup_u (\mu_A(u) \wedge \mu_{Ai}(u))$: matching coefficient

VERSION EP(1,1b) (DEMPSTER-SHAFER)

$$X \text{ isp } (p_1\backslash u_1 + \ldots + p_u\backslash u_n)$$

$$(X,Y) \text{ is } R$$

$$Y \text{ isp } (p_1\backslash R(u_1) + \ldots + p_n\backslash R(u_n))$$

Y is a fuzzy-set-valued random variable

$$\mu_{R(u_i)}(v) = \mu_R(u_i, v)$$

VERSION GEP(0,0)

$$\frac{f(X) \text{ is } A}{g(X) \text{ is } g(f^{-1}(A))}$$

$$\mu_{g(f^{-1}(A))}(v) = \sup_u(\mu_A(f(u)))$$

subject to

$$v = g(u)$$

GENERALIZED EXTENSION PRINCIPLE

$f(X)$ is A

$g(Y)$ is B

$Z = h(X,Y)$

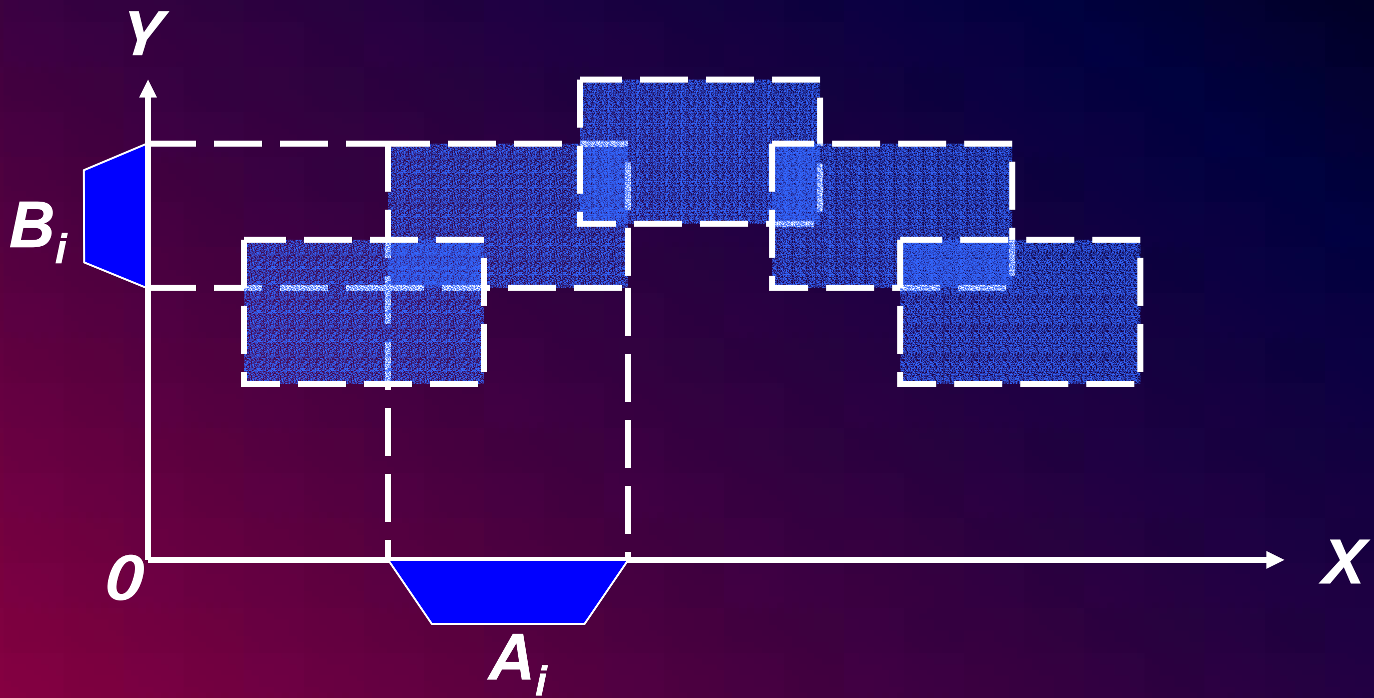
Z is $h\ (f^{-1}(A),\ g^{-1}\ (B))$

$$\mu_Z(w) = sup_{u,v}\ (\mu_A(f(u)) \wedge \mu_B(g(u)))$$

subject to

$$w = h(u,v)$$

U-QUALIFIED EXTENSION PRINCIPLE



If X is A_i then Y isu B_i, $i = 1, \ldots, n$

X isu A

Y isu $\Sigma_i \, m_i \wedge B_i$

$m = \sup_u (\mu_A(u) \wedge \mu_{A_i}(u))$: matching coefficient

THE ROBERT EXAMPLE

LAZ 4/24/2003

THE ROBERT EXAMPLE

- *the Robert example relates to everyday commonsense reasoning– a kind of reasoning which is preponderantly perception-based*

- *the Robert example is intended to serve as a test of the deductive capability of a reasoning system to operate on perception-based information*

THE ROBERT EXAMPLE

- the Robert example is a sequence of versions of increasing complexity in which what varies is the initial data-set (IDS)

version 1

IDS: usually Robert returns from work at about 6 pm

questions:

q_1 : what is the probability that Robert is home at t^* (about t pm)?

q_2 : what is the earliest time at which the probability that Robert is home is high?

CONTINUED

version 2:

IDS: usually Robert leaves office at about 5:30pm, and usually it takes about 30min to get home

q_1, q_2 : same as in version 1

version 3: this version is similar to version 2 except that travel time depends on the time of departure from office.

q_1, q_2: same as version 1

THE ROBERT EXAMPLE (VERSION 3)

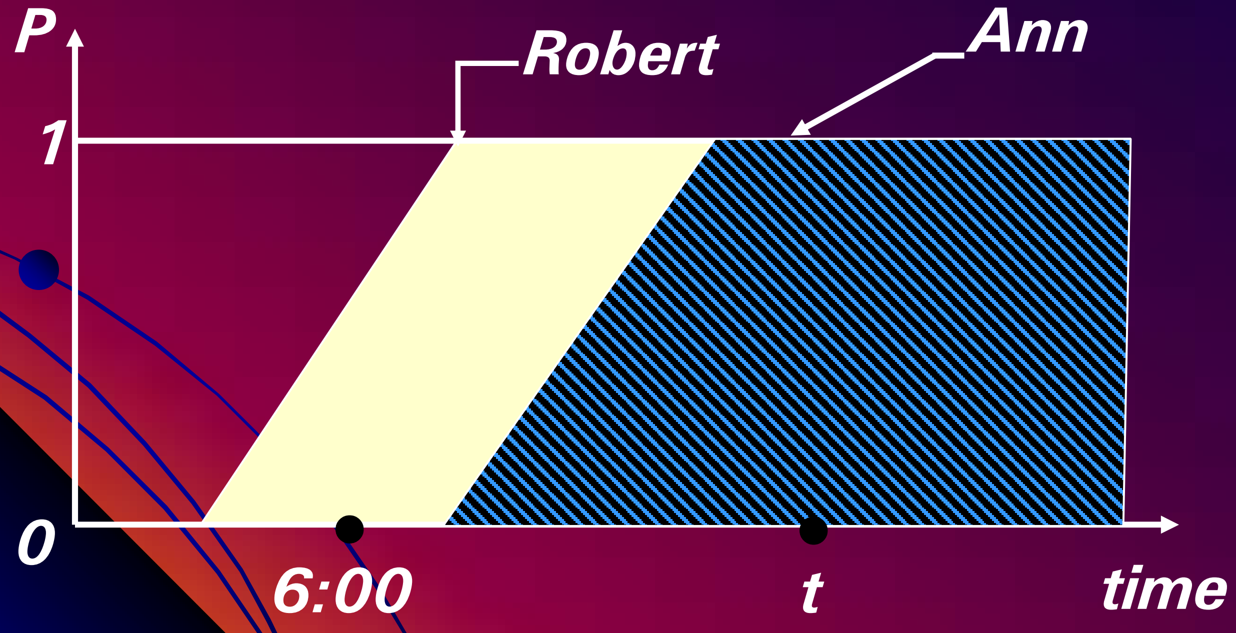
IDS: Robert leaves office between 5:15pm and 5:45pm. When the time of departure is about 5:20pm, the travel time is usually about 20min; when the time of departure is about 5:30pm, the travel time is usually about 30min; when the time of departure is about 5:40pm, the travel time is about 20min

- *usually Robert leaves office at about 5:30pm*
- *What is the probability that Robert is home at about t pm?*

THE ROBERT EXAMPLE

Version 4

- *Usually Robert returns from work at about 6 pm*
  Usually Ann returns from work about half-an-hour later
  What is the probability that both Robert and Ann are home at about t pm?



THE ROBERT EXAMPLE

Version 1.

My perception is that Robert usually returns from work at about 6:00pm

q_1 : What is the probability that Robert is home at about t pm?

q_2 : What is the earliest time at which the probability that Robert is home is high?

PROTOFORMAL DEDUCTION

THE ROBERT EXAMPLE

IDS p: usually Robert returns from work at about 6 pm.
TDS q: what is the probability that Robert is home at about t pm?

1. precisiation:
   p ⟶ Prob {Time (Robert returns from work is about 6 pm} is usually
   q ⟶ Prob {Time (Robert is home) is about t pm} is ?D
2. calibration: $\mu_{usually}$, μ_{t^*}, t^* = about t
3. abstraction:
   p* ⟶ Prob {X is A} is B
   q* ⟶ Prob {Y is C} is ?D

CONTINUED

4. search in Probability module for applicable rules (top-level agent)

$$\frac{\text{Prob } \{X \text{ is } A\} \text{ is } B}{\text{Prob } \{Y \text{ is } C\} \text{ is } D}$$ not found

found:

$$\frac{\text{Prob } \{X \text{ is } A\} \text{ is } B}{\text{Prob } \{X \text{ is } C\} \text{ is } D}$$

$$\frac{\text{Prob } \{X \text{ is } A\} \text{ is } B}{\text{Prob } \{f(X) \text{ is } C\} \text{ is } D}$$

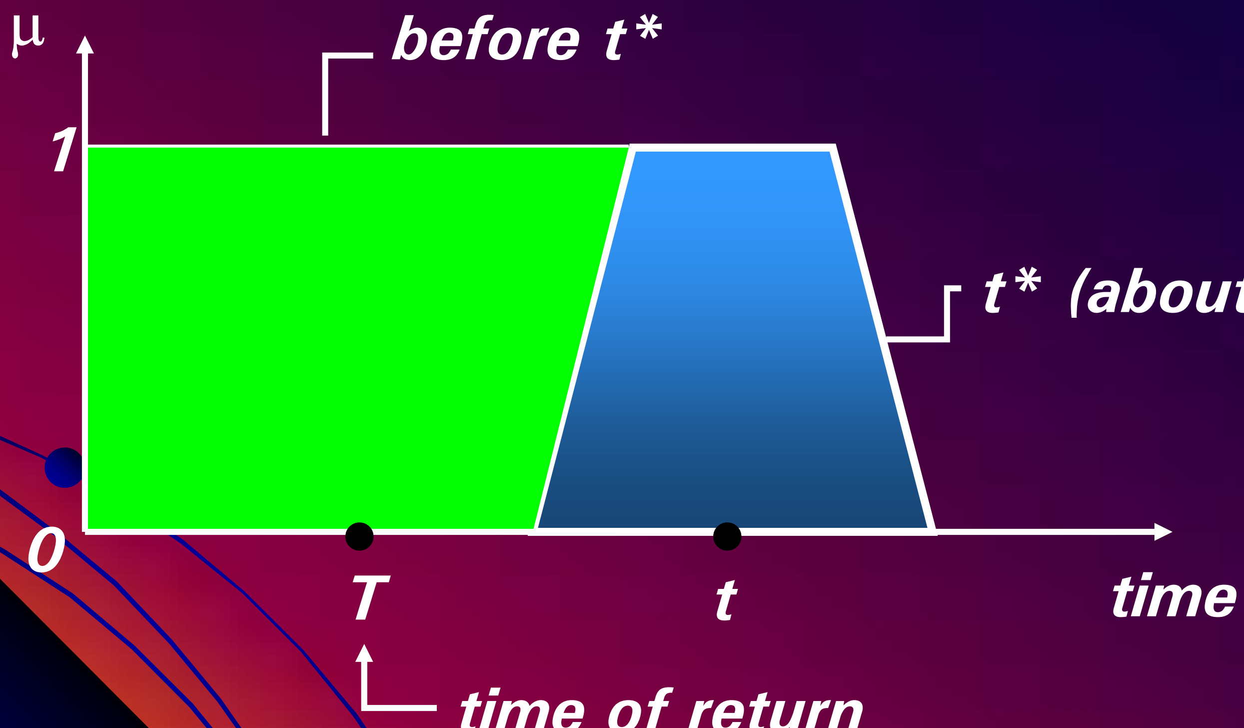
5. back to IDS and TDS. Go to WKDB (top-level agent)

- A/person is at home at time t if A returns before t
- Robert is home at t* =Robert returns from work before t*

THE ROBERT EXAMPLE

event equivalence

Robert is home at about t pm = Robert returns from work before about t pm



Before about t pm = $\leq_{o}$ *about t pm*

CONTINUED

6. back to Probability module

$$\frac{\text{Prob }\{X \text{ is } A\} \text{ is } B}{\text{Prob }\{X \text{ is } C\} \text{ is } D}$$

$$\mu_D(v) = sup_g(\mu_B(\int_U \mu_A(u)g(u)du))$$

$$v = \int_U \mu_c(u)g(u)du$$

7. Instantiation : D =Prob {Robert is home at about 6:15}
X =Time (Robert returns from work)
A = 6*
B =usually
C = $\leq$ 6:15*

THE BALLS-IN-BOX EXAMPLE

- a box contains N balls of various sizes
- my perceptions are:
  - a few are small
  - most are medium
  - a few are large

  IDS (initial data set)
- a ball is drawn at random
- what is the probability that the ball is neither small nor large

PERCEPTION-BASED ANALYSIS

a few are small $\longrightarrow \frac{1}{N} \Sigma$ *Count(small) is few* $\longrightarrow Q_1$ *A's are B's*

most are medium $\longrightarrow \frac{1}{N} \Sigma$ *Count(medium) is most* $\longrightarrow Q_2$ *A's are C's*

a few are large $\longrightarrow \frac{1}{N} \Sigma$ *Count(large) is few* $\longrightarrow Q_3$ *A's are D's*

$A = \{u_1,\ldots,u_n\}$; u_i =size of i th ball; u= $(u_i,\ldots,u_n)$

$\Pi_1(u_1,\ldots,u_n)$: possibility distribution function of $(u_i,\ldots,u_n)$ induced by the protoform Q_1 A's are B's

$$\Pi_1(u_1,\ldots,u_n) - \mu_{Q_1}\left(\frac{1}{N}\Sigma_i \mu_B(u_i)\right)$$

CONTINUED

$\Pi(u_1,\ldots,u_n)$: *possibility distribution function induced by IDS*

$$\Pi(u_1,\ldots,u_n) = \Pi_1(u_1,\ldots,u_n) \wedge \Pi_2(u_1,\ldots,u_n) \wedge \Pi_3(u_1,\ldots,u_n)$$

query: (proportion of balls which are neither large nor small) is? Q_4

$$Q_4 = \frac{1}{N} \Sigma_i((1-\mu_{small}(u_i) \wedge (1-\mu_{large}(\mu_i))$$

protoformal deduction rule (extension principle)

$$\mu_{Q_4}(v) = sup_u(\Pi_1(u) \wedge \Pi_2(u) \wedge \Pi_3(u))$$

subject to $V = \frac{1}{N} \Sigma_i((1-\mu_{B_1}(\mu_i)) \wedge (1-\mu_{B_3}(u_i)))$

SUMMATION—BASIC POINTS

- *Among a large number and variety of perceptions in human cognition, there are three that stand out in importance*
  1. *perception of likelihood*
  2. *perception of truth (similarity, compatibility, correspondence)*
  3. *perception of possibility (ease of attainment)*
- *These perceptions, as most others, are a matter of degree*
- *In bivalent-logic-based probability theory, PT, only perception of likelihood is a matter of degree*
- *In perception-based probability theory, PTp, in addition to the perception of likelihood, perceptions of truth and possibility are, or are allowed to be, a matter of degree*

CONCLUSION

- *Conceptually, computationally and mathematically, perception-based probability theory is significantly more complex than measurement-based probability theory.*
- *Complexity is the price that has to be paid to reduce the gap between theory and reality.*

COMMENTS

from preface to the Special Issue on Imprecise Probabilities, Journal of Statistical Planning and Inference, Vol. 105, 2002

"There is a wide range of views concerning the sources and significance of imprecision. This ranges from de Finetti's view, that imprecision arises merely from incomplete elicitation of subjective probabilities, to Zadeh's view, that most of the information relevant to probabilistic analysis is intrinsically imprecise, and that there is imprecision and fuzziness not only in probabilities, but also in events, relations and properties such as independence. The research program outlined by Zadeh is a more radical departure from standard probability theory than the other approaches in this volume." (Jean-Marc Bernard)

CONTINUED

From: Peter Walley (Co-editor of special issue)

"I think that your ideas on perception-based probability are exciting and I hope that they will be published in probability and statistics journals where they will be widely read. I think that there is an urgent need for a new, more innovative and more eclectic, journal in the area. The established journals are just not receptive to new ideas - their editors are convinced that all the fundamental ideas of probability were established by Kolmogorov and Bayes, and that it only remains to develop them! "

CONTINUED

From: Patrick Suppes (Stanford)

"I am not suggesting I fully understand what the final outcome of this direction of work will be, but I am confident that the vigor of the debate, and even more the depth of the new applications of fuzzy logic, constitute a genuinely new turn in the long history of concepts and theories for dealing with uncertainty."

STATISTICS

Count of papers containing the word "fuzzy" in the title, as cited in INSPEC and MATH.SCI.NET databases. (data for 2002 are not complete)

Compiled by Camille Wanat, Head, Engineering Library, UC Berkeley, April 17, 2003

	INSPEC/fuzzy	Math.Sci.Net/fuzzy
1970-1979	569	443
1980-1989	2,404	2,466
1990-1999	23,207	5,472
2000-present	8,745	2,319
1970-present	34,925	10,700